GROUPING OF THE FIFTY STATES: A COMPARISON OF FACTOR ANALYTIC METHODS

By

MARTHA JEAN CHANG

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1971

ACKNOWLEDGEMENTS

I acknowledge with gratitude the help of a number of individuals whose support made my graduate work and this dissertation possible.

Dr. Wilson H. Guertin, as my adviser and chairman of my supervisory committee, has been all that an adviser should be. He has in turn listened, prodded, encouraged, scolded, and praised. It was from him that I learned the intricacies of factor analysis.

Dr. Alfred E. Clubok and Dr. Willias B. Ware have been valuable members of my supervisory committee because they have been able to see my research from such different viewpoints.

Many other professors have been encouraging and helpful during my graduate years at Florida. Among them I mention Dr. Hal G. Lewis, Dr. Vynce A. Hines, and Dr. Robert S. Soar.

My appreciation goes especially to my very patient family--to Penny and Perry, who have tolerated my months of preoccupation, and to Richard, who because he once traveled this road himself, has gone beyond support to active involvement.

TABLE OF CONTENTS

LIST OF TABLES

TABLE		PAGE

TABLE		PAGE
37	Shape Family 2, Pattern 1	83
38	Shape Family 2, Pattern 2	84
39	Shape Family 2, Pattern 3	85
40	Shape Family 3, Pattern 1	87
41	Shape Family 3, Pattern 2	88
42	Shape Family 3, Pattern 3	89
43	Shape Family 4, Pattern 1	91
44	Shape Family 5, Pattern 1	94
45	Shape Family 6, Pattern 1	95
46	Unclassified States	96
A.1	Varimax Rotated Factor Loadings: R Analysis .	111
A.2	Simple Loadings Factors: Q Analysis Based on Raw Score Distances	114
A.3	Maxplane Factors: Q Analysis Based on Raw Score Distances	116
A.4	Oblimin Factors: Q Analysis Based on Raw Score Distance	118
A.5	Simple Loadings Factors: Q Analysis Based on Unit Weighted Factor Score Distances	120
A.6	Simple Loadings Factors: Q Analysis Based on Variance Weighted Factor Score Distance . .	122
A.7	Simple Loadings Factors: D Analysis Based on Factor Scores Weighted According to Percentage of Educational Variables	124
A.8	Varimax Rotated Factor Loadings: Q Analysis Based on Correlation	126

LIST OF FIGURES

FIGURE PAGE

Abstract of Dissertation Presented to the Graduate Council of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

GROUPING OF THE FIFTY STATES:
A COMPARISON OF FACTOR ANALYTIC METHODS

By

Martha Jean Chang

December, 1971

Chairman: Wilson H. Guertin
Major Department: Foundations of Education

Factor analysis has made important contributions in such areas of education as educational psychology and test construction, but it has rarely been used in social foundations of education. Moreover, it has been used mainly to determine factors or dimensions within test items or personality measures. This study focuses upon the comparative grouping of cases, specifically, the grouping of the 50 states on the basis of 53 socioeconomic, political, and educational characteristics.

The study attempts to answer the following questions: Which is preferable for grouping political units, correlation indices or distance indices? What grouping differences result in distance analysis by using standardized raw scores as opposed to factor scores? How does giving factor scores different weights affect the groupings? How useful is profile analysis in determining types of political units?

A factor analysis of the variables determined the following five interpretable factors: Socioeconomic, Racial-Educational,

Educational Expenses vs. Manufacturing, Frontier, and Agriculture. A comparison of factor analytic groupings on correlation indices and distance indices showed that the use of distance measures resulted in clearer, more inclusive groupings. Conceptually, research focusing on the grouping of political units is usually more interested in grouping on distance, or score magnitude, than on profile shape, as measured by correlation. However, the correlation approach is useful as employed in profile analysis, which proved valuable in the interpretation of the groups in terms of variable differences.

Four distance analyses were made, one using standardized raw scores and three employing weighted factor scores, specifically, unit weights, weights based on the percentage of variance accounted for by each factor, and weights based on the percentage of educational variables loading on each factor. The resulting groupings were not greatly different from each other. The results of a study of Florida counties, presented as a contrast to the 50 states study, however, showed that using unit weighted factor scores resulted in quite different groupings than using standardized raw scores. The conclusion is that the use of standardized data variables in distance analysis is usually preferable because of the information retained on the cases' scores on individual variables. Where factor scores are used, comparisons should be run to explore possible differences resulting from the use of different weights.

CHAPTER 1

INTRODUCTION

When confronted with a number of entities, man almost invariably groups. The limits of his own mind make it impossible for him to wrench understanding from consideration of each individual entity. Only as all or most of the entities are examined, analyzed, and categorized can man begin to make sense out of the domain in which he works. This categorizing is usually made on the basis of subjective impressions, which may or may not be effective. Out of informal grouping experiences developed more precise methods of grouping based on quantified data, that is, methods utilizing measurement of some relevant aspect of the domain. Using a single measurement one could break down the entities into such categories as high, medium, and low, or even rank each entity. This was fine as far as it could go, but only as man could take into account multidimensional considerations could there be truly effective grouping.

Factor analysis offers a general empirical method for grouping any given body of entities on the basis of a number of characteristics. But factor analysis is fraught with a myriad of decisions: what variables? principal components or principal axes? which rotation? how many factors? use of raw data or factor scores? which grouping method?

One general purpose of this study is to provide a case history of these decision-making processes. Although each question is applied specifically to this case, the decisions reached can have applicability to a wide range of domains.

This study is designed to compare methods of grouping cases. In the factor analytic application of grouping of cases by social scientists several methods have been used. The choice of method seems to be tied closely to the discipline; as a result, geographers, for instance, may use one method and political scientists another, even though there is no apparent clear-cut reason other than convention. Profile analysis (Gaertin, 1966) has evidently not been used at all by the social scientists; correlation measures have been used when distance would sometimes be preferable; and the use of distance analysis has been inadequately explored. There is a need for studies which focus upon a comparison of grouping methods rather than upon the use of any particular method. Russett's 1967 study, using nation-state data, went far in this direction, but much more needs to be done. He stated in comparing alternate grouping procedures: "It does require that, to settle the question finally, we have a large sample of rigorous comparisons in which all the comparisons are reported, not just those that produce very similar (or different) results as the analyst happens to want to show" (p. 57).

The present study aims at providing just such a sample. The focus will be on the grouping of socioeconomic, political,

and educational characteristics of the 50 states of the United States and, more importantly, on grouping the states themselves on the basis of these same characteristics. In the grouping of the states, several methods will be used, analyzed, and reported.

The framework for this study is the socioeconomic foundations of education. The technique of factor analysis has made important contributions in education in such areas as educational psychology and test construction. In contrast to what has been taking place during the past two decades in the social sciences in general, however, there has been little carry-over into areas where education and the social sciences meet--the historical and socioeconomic foundations of education. This study will employ some of the variables commonly used in this area of education but, focusing on some grouping methods, it will not attempt to apply the findings to educational theory. Further studies are needed which will firmly tie factor analysis and grouping methods in general to substantive areas of social foundations of education.

Review of Literature

Factor analysis has long been an important tool of the behavioral scientists and as such has done much to develop scientific theories in such areas as personality and intelligence. Scholars in the social sciences were, however, much slower to see the possibilities of this mathematical technique

in their own work. Indeed, some of the early statistical studies of political units were made by psychologists rather than social scientists. An early example of a psychologist applying statistical techniques to the study of political units was E. L. Thorndike's Your City (1939). Much more sophisticated work was done by psychologist R. B. Cattell and his associates on factor analytic studies of characteristics of nations (1949, 1950, 1951b). In the meanwhile political scientists were beginning to apply factor analysis to voting records (Gosnell and Gill, 1935; Gosnell and Schmidt, 1936; Gosnell, 1937; Carlson and Harrell, 1942; Harris, 1948) and characteristics of political units (Ogburn, 1935; Price, 1942). Since those early years factor analysis has become increasingly important to areas of sociology, political science, anthropology, international relations, and geography.

Most of the factor analytic studies carried on in these disciplines, except perhaps geography, have been in the grouping of variables (characteristics). However, interest has been growing in the grouping of political units (cases) (Cattell, 1950; Hatt, Farr, and Weinstein, 1955; Banks and Gregg, 1965; Russett, 1967). This grouping has been closely tied to an area which is of particular concern in geography, political science, and sociology--regionalization. In the 1930s much work was done on regionalism in the United States by H. W. Odum and associates (examples, Odum, 1936; Odum and Moore, 1938), though without the aid of factor analysis.

In the 1960s geographers who studied regionalism employed the factor analytic and related statistical methods used by B. J. L. Berry and associates (see Berry, 1961) in what geographers are now beginning to speak of as a Quantitative Revolution (Smith, 1971, p. 10). Political scientists also became heavily involved in factor analytic studies of political regions. Two prominent examples were B. M. Russett's study of international regions (1967) and R. J. Rummel's Dimensionality of Nations Project.

Despite these far-reaching studies many gaps remain. Few studies have been made grouping characteristics using as cases the states of the United States (Hofstaetter, 1951; Schutz, 1956; Hofferbert, 1968; Cole, 1967), and even fewer in the area of grouping the states (Sutton, 1967). Again, only two factor analytic works can be cited in the area of education in the United States (Hofstaetter, 1951; Schutz, 1956), both done before computers were available generally. Both of these studies focused on the grouping of characteristics. Schutz made no attempt to delineate groups; Hofstaetter made a rudimentary attempt in the form of graphing the factor scores of six of the states. This present study, in addition to presenting a comparison of methods of grouping political units, aims at providing an example of relating a grouping of socioeconomic, political, and educational characteristics of the states to groupings of the states themselves on this same data.

CHAPTER 2

PROCEDURE

Factor analysis is a procedure whereby one may examine a matrix of numbers to see which of the entities being studied group together because of a tendency to vary in the same way. The matrix to be analyzed is a slice of the "data cube," a concept originally developed fully by Cattell (1952; see also Cattell, 1966, and Rummel, 1970, pp. 192-202). The focus is on two dimensions at some point on the third dimension. Thus one can study a cases by characteristics slice for one occasion, an occasions by characteristics slice for one case, or a cases by occasions slice for one characteristic. The focus for this study will be on the cross-sectional slice consisting of the 50 states (cases) by socioeconomic, political, and educational variables (characteristics) for the year 1960 (occasion), the last year for which extensive data are available.

Once the data slice is determined one still has several possibilities for focusing the analysis. If factor analysis can be viewed as a formal grouping procedure, one can see that, given a cases by characteristics matrix, he may group either cases or characteristics. Historically, grouping of characteristics came first in the pioneering work of Spearman (1904) and Thurstone (1931). As more experience

was gained in the use of factor analysis, some psychologists began grouping cases (Stephenson, 1936; Burt, 1937). This type of analysis Burt termed transpose factor analysis. Cattell (1952), in developing the full data cube, designated the analysis of the interrelationship of characteristics R technique, and the analysis of cases, Q technique. These terms have become widely accepted and they, and the phrases R analysis and Q analysis, with identical meaning, will be used throughout this study.

This study will first present an R analysis, in which the focus will be on the dimensions or factors of the characteristics of the states. The results of several methods of transposing the data and studying the interrelationship of the states will then be reported. The latter type of analysis will be the heart of the study, since this is the area in which so little work has been done. A more detailed description of the comparative methods for this Q analysis will be considered shortly.

The computer programs used for the analyses were those developed for the Educational Evaluation Library, the University of Florida. These programs are described in Guertin and Bailey (1970, pp. 293-314). The analyses employ the principal axes method, with the multiple R, where an inverse of the correlation matrix is possible, used as the original communality estimate. Where an inverse is not possible, the highest correlation coefficient is used in place of the multiple R. The computer programs provide for iteration

to reestiaate communalities on the basis of the raw sums of squared loadings. After the number of factors to be rotated was determined in order to produce the clearest picture of common factor space (Guertin and Bailey, 1970, pp. 109-20), the factors were rotated both orthogonally, according to the Varimax criterion (Kaiser, 1958), and obliquely. Computer programs used for oblique rotations were Simple Loadings (Jennrich and Sampson, 1966), Oblimin (Carroll, 1958), and Maxplane (Cattell and Muerle, 1960; Eber, 1966). In all cases factors were rotated by the Simple Loadings procedure, and the results compared to that of the Varimax rotation to see which rotation best approximated simple structure. Simple structure, in short, means that each variable or case, depending upon which is being grouped, can be explained by as few factors as possible. If the results indicated that the factors were essentially independent, the orthogonal rotation was accepted. If the oblique solution resulted in better structure, with the factors clearly intercorrelated, rotations were sometimes run on all three oblique rotation procedures and the rotations were then compared. In all cases the results of the Simple Loadings procedure seemed superior. (For a comparison of these oblique procedures, see Bailey and Guertin, 1970.) The program available for calculating factor scores, or more accurately, factor score estimates, utilized the complete method based upon the inverse of the correlation matrix (Harman, 1967, pp. 350-60).

Methods for Grouping Cases

As outlined above, an R-type factor analysis correlates and groups characteristics into factors or dimensions. The same data can be transposed so it will group cases. Such grouping may be obtained by two methods. One is very close to the factor analytic methods used for the R-type analysis. Scores on the variables are first standardized to uniform means and standard deviations and then intercorrelated with scores of other cases. The principal axes method is then used, as in R-type analysis, to determine dimensions or factors within the correlation matrix. In order to understand clearly what the resulting factor structure means, one must bear in mind what correlation itself represents. A simple correlation between two cases is an index of relation between them in variance overlap terms, i.e., a perfect correlation of 1.0 between the profiles of the cases means that as one looks across variables, whenever the score of one case moves in one direction, the score of the other will always move in the same direction. Thus two states with very different magnitudes in socioeconomic measures could correlate appreciably if in each case the same variables changed together in the same--or opposite, if negatively correlated--direction. A simplified example would be that two states might be highly correlated if for each the average salary of teachers had a near constant relationship with the percentages of population urban and percentages who finished college, irrespective of the magnitude

of these [illegible]. Figure 1 presents this example in graphic form.

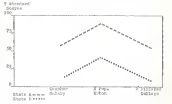

Fig. 1 Example of High Correlation

Q-type analysis, based on correlation, therefore, groups cases on the basis of the shape of their profiles while ignoring level. (For further discussion, see Guertin and Bailey, 1970, pp. 264-66.) For many analyses such a discrimination is all that is wanted, for the attempt is to delineate certain shapes. In grouping states, however, one is usually interested in discovering similarities not only in profiles but in magnitude, in bringing together clusters of cases whose shapes are similar and who are found at the same level. This important, but little known approach, is that taken by profile analysis, which compares shapes and then clusters together cases with

similar shapes at similar levels (Guertin, 1966). For this present study the data were transposed on the basis of the intercorrelation matrix (Q based on r) and the profile analysis approach was then used as a further refinement.

Another method for transposing data, in contrast to the above relational approach, is one based upon distance measures. The distance measure which was used for this study was introduced by Cronbach and Gleser (1953). This index, known as d, is simply the square root of the sums of the squared distances between two profiles across all variables. Arithmetic operations as developed by Guertin (1971) transfers this measure from a matrix of distance measures of dissimilarity to one of distance similarity indices (DSI); this matrix is subsequently factored as if it were a correlation matrix. The resulting factors bring together cases which are congruent, that is, similar in magnitude and shape. Figure 2 includes the highly correlated states introduced in Figure 1 and adds to them State C, with scores close to those of State A, but with a shape dissimilar to those of either of the other states. Q based on correlation would group State A and State B. Q based on distance would group State A and State C. These figures present in a nutshell the essential differences between the two approaches.

One question to be settled in using distance analysis is the type of data to be used. Guertin and Bailey (1970) suggest the use of the standardized data variables. Bummel

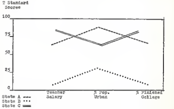

Fig. 2 Example of Correlation and Distance

(1970, p. 495) argues that the use of such data "artificially" weights correlated variables and, therefore, he prefers to use standardized factor scores obtained from an R-type analysis. For this study both methods were used for comparative purposes. In addition, a distance analysis was made using factor scores weighted in accordance with the percentage of total variance accounted for by each factor. A final distance analysis was made using factor scores weighted by the percentage of educational variables on each factor. Because of the relevance to this question, results from a study of Florida counties are presented which are quite different from the results of this present study (Chang, 1971).

Variables

If this were a study in which the emphasis was on the development of theory, the choice of variables (characteristics) would be of utmost importance. Instead the emphasis is upon technique or method and of importance only was the sampling of variables in the areas of socioeconomic, political, and educational characteristics. Variables were chosen from among those available which seemed capable of distinguishing between the states, on one hand, and of adequately measuring socioeconomic, political, and educational characteristics, on the other. One could, for instance, get clear groupings of states by the use of such variables as longitude and latitude, temperature, rainfall, and distance above sea level, but these geophysical measures were not what was wanted. (See Rummel, 1970, pp. 209-26, for a detailed discussion of suggestions for the selection of variables for factor analytic studies in the social sciences.) Guidance in the choice of variables came both from those found useful in other social science and educational research, and in the experience gained in two pilot studies, one on United States data and the other on Florida county data. For instance, it was found that by using both population size and total public school enrollment as marker variables, a size factor emerged. The decision had already been made not to force out a size variable. Therefore, only the total public school enrollment was retained as a marker variable. An attempt was made to sample variables in the following interrelated areas:

Economic characteristics
Population characteristics
Political characteristics

Crime indices
Population growth and migration
Educational characteristics
Educational support.

In the end, 53 measures were chosen; a list of the variables and sources will be found in the Appendix. All variables except one were used in the form of percentages and per capita. The exception is the variable total public school enrollment, already mentioned.

In working with the intercorrelation matrix on which factor analysis is usually based, some researchers feel it desirable to secure data on at least 100 replicates in order to reduce error variance (Guertin and Bailey, 1970, pp. 170f; Cattell, 1966, p. 236). Two issues are involved here. One is the use often made of factor analysis in taking a sample and generalizing the resulting factors to a population. In this study the emphasis is upon describing the factor structure of the characteristics rather than upon generalizing from a sample to a population. The states themselves represent the population. One cannot collect data on an N of 100 when the population is only 50.

The second issue is the reliability of the correlation coefficients of the entities themselves. One cannot have high reliability of intercorrelations without high reliability of data. Within the states, many of the data series are based upon observations of the entire population of each state and cannot be considered sample data to begin with. Where data are based on sampling, the sampling is random and much larger than commonly used. Cattell dealt with this

problem many years ago, and in defending his N of 69 nations, stated: "A population of 69 is small for the extraction of reliable factor loadings for many factors, but when, as in this case, it is the whole population, rather than a sample from a larger population, certain objections do not apply" (1949, pp. 44f).

One is justified in assuming a relatively high degree of reliability in the use of census data for several reasons. (a) Scale: The process of counting is a precise procedure, which is intrinsically more reliable than measuring continuous data. (b) Procedure: The data collection, though not under the control of the researcher, was thorough, based on long experience, and planned by statistical experts. The wording of questions has become increasingly skillful so that the data of recent census are sufficiently reliable. (See Folger and Nam, 1967, for a discussion of the reliability of census data as it applies to educational research.) (c) Sample size: This was discussed above. In summary, the data to be used in this study are highly reliable; therefore, one can be assured of analyzing mainly systematic variance as opposed to error variance.

A related problem is that of the relationship of the number of cases to the number of characteristics in the grouping of the latter. There is no agreement on any single ratio between the two except that it is, of course, desirable to have more cases than variables in R analysis. Some factor analysts favor a 2 to 1 lower bound, while others go as high

as 5 to 1 (Cattell, 1966). Here evaluation of Cattell's changing view on this problem is relevant and instructive. He once suggested a 4 to 1 ratio (Cattell, 1952). Subsequently, however, he asserted: "Admittedly, a ratio of at least 2 to 1 is helpful as a rough guide; but closer examination indicates ways in which exceptions are permissible under pressure of circumstances" (Cattell, 1966). It would seem that this more tolerant view emanated from his own research experience. In one of his factor analytic studies, 72 characteristics are factor analyzed over 69 nations (1949); in another, the same 72 are analyzed over 40 nations (1951). Notice that in both studies the number of cases does not, as it is supposed to, exceed that of variables; instead, the reverse is true. Yet his resulting factor structures proved to be substantively meaningful. It is evident, then, that, while trying to make the ratio as large as possible in data gathering, one must not give up a research project simply because the ratio in question happens to be less than 2 to 1. The ultimate test for adequacy of the ratio is whether the most stable factor structure, obtained from the data based on a less than desirable ratio, makes sense from the standpoint of substance.

In a study such as this one where the plan is to group both cases and characteristics, the problem is further complicated since one would like to have an N of 100 both ways--100 cases for R analysis and 100 characteristics for Q analysis. Even if this were possible, one could not at the same time satisfy the 2-to-1 rule, or any other desirable ratio. One might conceivably drop variables in the final grouping of

characteristics, but he would first have to prove that the factor structures of both samples of variables were similar. This study therefore employs a 50 by 53 matrix, roughly a square matrix, for both analyses. With all of the complexities involved, the use of reliable, widely ranging, meaningful variables weighs more heavily in decision-making than mere algebraic considerations.

One additional reason for having more cases than variables is that this is necessary in order to obtain an inverse of the correlation matrix. The procedures used are dependent upon the inverse at two points. One is in setting a communality estimate and the other is in calculating factor scores, both mandatory to this study. Fortunately, substitute techniques are available or have been developed at both points. The computer programs themselves provide a substitute for the first for, although they first attempt to use multiple R, dependent on the inverse, for the original communality estimate, they provide for the use of highest r if this is not attainable. The method which has been devised for getting around the indeterminacy of the inverse of r in calculating factor scores will be explained in Chapter 4.

To summarize, the results and discussion of these comparative factor analytic techniques will be divided into three sections. Chapter 3 will describe the R analysis. The four Q analyses based on distance will be compared in Chapter 4. Finally, Chapter 5 will present Q analysis based on correlation and the related profile analysis.

CHAPTER 3

R ANALYSIS

As a preliminary to transposing the data, an R analysis was made, correlating and grouping the 53 variables. The principal axes method was used with the highest coefficients employed as the original communality estimates. Three iterations brought satisfactory convergence between the final communality estimates and the row sums of squared loadings. From this, 18 factors, accounting for 96 percent of the total score variance, were extracted. On the basis of the size of their latent roots, the intercorrelations, and a comparison of factor structure stability of Varimax rotations of several alternate number of factors, an eight-factor solution was selected. These eight factors account for 85 percent of the total score variance, 89 percent of the common variance. A check on an oblique solution, specifically, Simple Loadings (Jennrich and Sampson, 1966), indicated, as expected, that the factors were essentially independent. Hence, the orthogonal Varimax solution was chosen. Of the eight factors rotated, five could be interpreted meaningfully. These five factors, presented in Table 1, account for 77 percent of the total score variance.

A table of the complete Varimax solution will be found in the Appendix. Tables of individual factors in this chapter

TABLE 1
Socioeconomic, Political, and Educational Dimensions of the Fifty States

Factor	% of Total Score Variance
Socioeconomic	21.6
Racial-Educational	18.8
Educational Expense vs. Manufacturing	13.6
Frontier	11.7
Agriculture	8.5

and the two that follow include only the highest loaded variables. Following the practice of Cattell, the limit for "highest loaded variables" is drawn differently in different factors according to the natural division points and the number of variables with sizeable loadings on each factor (Cattell and Adelson, 1951, p. 196). For instance, all of the variables listed on Table 2 (Factor 1) have loadings of ±.45 or above; those on Table 8 (Factor 5), ±.30. The purpose of these summary tables is to include sufficient variables to explain the substantive meaning of the factors.

Factor 1, labeled Socioeconomic, is comparable to similar factors in other factor analytic studies of political units. It exhausts 22 per cent of the total factor variance, and is thereby the weightiest factor. As shown on Table 2, 20 of the variables have loadings of ±.50 or above on this factor. The states which contribute most variance to this factor tend to be urban states which have a relatively high socioeconomic level, as evidenced by a number of highly loaded economic variables and such educationally related variables as percentage

TABLE 2
Varimax Rotated Factor Loadings
R Analysis
Factor 1: Socioeconomic

Characteristic	Factor Loading
Average teacher salary	.85
% population urban	.85
Per capita personal income	.79
% families with annual income $10,000 or more	.78
% employed persons in white-collar jobs	.77
Per pupil expenditures	.77
% population rural nonfarm	-.77
Per capita value selected services	.75
Personal income per pupil	.71
% families with annual income less than $3000	-.69
% population 25 years and over who completed college	.69
% 5 and 6 year olds enrolled	.66
% population foreign born	.63
% population 25 years and elder who completed high school or more	.57
% population rural farm	-.56
Per capita value retail sales	.55
High school graduates in 1959-60 as % of 8th grade enrollment in 1955-56	.54
Per capita burglary	.53
% change in public school enrollment 1950-51 to 1960-61	.51
% population increase, 1950-60	.50
Per capita state and local expenditures for education	.49
Total public school enrollment	.48
% of children in elementary school enrolled in private school	.48
% employed persons in agriculture	-.47
% total population enrolled in public school	-.45

of employed persons in white-collar jobs (.77), per pupil expenditures (.77), and percentage of the over 25-year-old population who has completed college (.69). The growth variables also have moderate positive loadings here.

Factor 2, shown in Table 3, is a bipolar factor labeled Racial-Educational. Percentages of the population White (.76) and Black (-.89) and most of the variables measuring

TABLE 3
Varimax Rotated Factor Loadings
R Analysis
Factor 2: Racial-Educational

Characteristic	Factor Loading
% population 21 and over voting in presidential election	.91
% population Negro	-.89
% population 25 years and over with less than 5 years of schooling	-.89
Per capita murder	-.87
% registrants failed mental test of Selective Service	-.85
Median school years nonwhite as % of median school years White	.82
% population White	.76
% of revenue for public elementary & secondary education from local sources	.64
Per capita value retail sales	.62
Pupils per classroom teacher	-.58
% families with annual income less than $3000	-.58
High school graduates in 1959-60 as % of 8th grade enrollment in 1955-56	.57
% population 25 years and older who completed high school or more	.55
% population 14-17 enrolled in school	.54
% population 65 years of age and over	.52
% of children in elementary school enrolled in private school	.43
% of 5 and 6 year olds enrolled	.41

educational characteristics of the population load highly on this factor, showing high availabilities of educational opportunities for the White culture, contrasted to lack of such opportunities for the Black. Thus this is both an educational-cultural factor and a bipolar racial factor. The political variable, percentage of the over-21 population who voted in the presidential election, loads highly here (.91). The data do not reveal whether this variable is more closely connected with the racial aspect or the educational

aspect of this factor. Instead the intercorrelation matrix on which this factor analysis is based indicates that voting participation in the 1960 presidential election had a relationship to both aspects. Table 4 gives some of the variables

TABLE 4
Percentage Adults Voting in the Presidential Election Correlation with Selected Variables

Variables	Correlation
Per capita murder	-.87
% population Negro	-.86
% population 25 years and over with less than 5 years of schooling	-.86
% registrants failed mental test of Selective Service	-.85
Median school years nonwhite as % of median school years white	.75
% population white	.73
Per capita value retail sales	.69
% population 25 years and older who completed high school or more	.68
% families with annual income less than $3000	-.67

which have a high correlation with this variable. Per capita murder also loads highly on Factor 2. Hackney (1969) gives a quantitative-historical explanation for the tie between this variable and the percentage of the population Negro.

Variables indicating high socioeconomic level do not load on this factor, for the variance of most of these has been exhausted by Factor 1. Instead this factor has a distinctive low socioeconomic character. This becomes evident in studying the rankings of the factor scores of the 50 states. A successive ranking program was used to rank each state on factor scores on each of the five factors. Table 5 shows

TABLE 5
Rankings of Selected States on Factors 1 and 2

Rank on Factor 1		Rank on Factor 2
49	New Hampshire	1
48	Maine	2
46	Vermont	3
42	Idaho	4
32	Iowa	5
31	Nebraska	6
.	.	.
47	Arkansas	44
44	North Carolina	45
36	Georgia	46
38	Alabama	47
45	South Carolina	48
20	Louisiana	49
41	Mississippi	50

that states ranking at both extremes of Factor 2 are states which have low ranks on Factor 1. Thus states which contribute variance to Factor 2 tend to be states which are low in Socioeconomic. Finally, by way of comparison, Factor 2 is related to Schutz' factor, Intellectual Climate (1956) and to Hofstaetter's Absence of Racial Discrimination (1951).

The variables loading highly on Factor 3, Educational Expenses vs. Manufacturing, may be seen in Table 6. This is another bipolar factor with such variables as a high percentage of personal income spent for education, a high percentage of persons employed as teachers, and a high percentage of school-age population on one end, and high per capita value added by manufacturing and high percentage of persons employed by manufacturing on the other. The reason for the inverse relationship between these two is not evident from the factor.

TABLE 6
Varimax Rotated Factor Loadings
R Analysis
Factor 3: Educational Expenses vs. Manufacturing

Characteristics	Factor Loading
% of state and local government employees who are teachers	.81
Expenditures for public elementary and secondary education as % of personal income	.81
% employed persons in manufacturing	-.74
% employed persons in educational services	.71
Per capita value added by manufacturing	-.71
Per capita motor vehicle registrations	.67
% total population enrolled in public school	.59
Per capita state & local expenditures for education	.58
Per capita value farm products	.49
% population 5 years and older migrant from a different county	.48
% of children in elementary school enrolled in private school	-.47
Birth rate	.46
% population 25 years and older who completed high school or more	.44
% population 14-17 enrolled in school	.44
Total public school enrollment	-.43
% civilian labor force male	.42
% employed persons in agriculture	.41
Personal income per pupil	-.40
Population density	-.37

The high expenditures for education are not related to such things as per pupil expenditures or teacher salaries, but instead are tied to the amount that each person in the state spends for education. Simply stated, a state which has a higher percentage of children and a smaller general population spread over a larger area will tend to spend more for education. This relationship may be seen in some of the variables loaded on this factor, variables such as percentage of the total population enrolled in public school (.59), the marker

variable indicating relative size of school-age population; birth rate (.46); total public school enrollment (-.43), the population size variable indicator; and population density (-.37).

Factor 4, labeled Frontier, appears on Table 7. Although it is perhaps an exaggeration to speak of the United States

TABLE 7
Varimax Rotated Factor Loadings
R Analysis
Factor 4: Frontier

Characteristic	Factor Loading
Average size school district	.92
Males per 100 females	.90
% population 5 years and older migrant from a different county	.68
% change in public school enrollment, 1950-51 to 1960-61	.61
Birth rate	.59
% population 65 years of age and over	-.58
% population increase, 1950-60	.55
% families with annual income $10,000 or more	.50
% population rural nonfarm	.46
% population other races	.46
% employed persons in manufacturing	-.36
% population 25 years and older who completed high school or more	.36
Median school years nonwhite as % of median school years White	.33
Total public school enrollment	-.32
Per capita value added by manufacturing	-.31
Per capita state & local expenditures for education	.31

having a land frontier, this is the predominant "mood" of this factor. The variables loading here measure such characteristics relating to developing geographic areas as a younger population, with a resulting higher birth rate, more males per 100 females than other areas, and comparatively little

manufacturing. Factor scores of the states show that Nevada, Hawaii, and particularly Alaska, contribute most of the variance to this factor. They are the only states which have factor scores on this factor greater than ±1.0, with Alaska having a standardized score of 6.0, Nevada 1.8, and Hawaii 1.5.

Factor 5, Agriculture (see Table 8), is straightforward and therefore easily interpreted. The highest loaded variables

TABLE 8
Varimax Rotated Factor Loadings
R Analysis
Factor 5: Agriculture

Characteristic	Factor Loading
Farm land as % of all land area	.80
Per capita value farm products	.70
% employed persons in agriculture	.70
% population rural farm	.69
% voted Democratic	-.54
Population density	-.41
% civilian labor force male	.40
% employed persons in manufacturing	-.36
% population 25 years and over who completed college	-.32
Average teacher salary	-.31
% population 14-17 enrolled in school	.31

are farm land as a percentage of all land area (.80), per capita value of farm products (.70), percentage of employed persons in agriculture (.70), and percentage of the population rural farm (.69). In keeping with the stereotype of a Republican farm belt, the percentage voting Democratic loads negatively (-.54).

The reduction of the 53 variables has produced the five interpretable factors analyzed. Of these five, Socio-economic, accounting for 21.6 percent of the total variance, loads the other four factors. It is also this factor which many earlier factor analytic studies have extracted. Similarly, the second factor, Social-Educational, finds its counterparts in earlier studies. Of particular interest in the results of the third factor is the finding that the population density is inversely associated with per capita educational expense. That is, the smaller the population density of a state, the greater its per capita educational expense. Alaska, Nevada, and Hawaii are largely responsible for extraction of the factor labeled "Frontier." The Republican farm belt, by and large, explains the emergence of the last factor, Agriculture.

These factors will be mentioned frequently in the next two chapters. With the discussion of the R analysis complete, we now turn to its transpose, Q analysis.

CHAPTER 4

Q ANALYSES BASED ON DISTANCE

When a Q-type analysis, grouping cases on the basis of characteristics, is made, the conventional approach is to factor analyze the intercorrelation matrix. Instead of introducing this oft-used type of analysis first and relating all other analyses to it, the choice has been made to introduce first the analysis which resulted in the clearest grouping for these data and then to relate Q based on correlation to it. This grouping method is a Q-type factor analysis based on a matrix of distance measures. This chapter will present a D analysis based on standardized raw data and three D analyses based on factor scores weighted, respectively, by unit weights, percentage of total score variance accounted for by each factor, and percentage of educational variables loading on each factor. It was hypothesized that the D analysis based on standardized raw scores, because it makes maximum use of the data, would produce the most discriminating grouping. This grouping was used as a criterion by which to judge the other D analyses. It was further hypothesized that the use of unweighted factor scores (in reality, factor scores employing unit weights), frequently used in social science research, would produce groupings which ignored the variance of cases on individual variables. Two states with similar factor scores might still

have important differences in scores on the variables which could help in discriminating between them. Furthermore, attaching equal importance to a factor which explains 22 percent of the total score variance, for instance, and a factor which explains 11 percent, could give distorted groupings. Two other D analyses were made with factor scores weighted according to two quite different criteria. The first criterion was the percentage of common variance accounted for by each factor. It was hypothesized that, of the three analyses employing factor scores, this particular grouping would most closely approximate that of the D analysis based on standardized raw data. Any difference between the groupings would be a result of loss of information. The final D analysis was made as an example of a second criterion by which factor scores could be weighted, by the percentage of variables present which represent the field of focus, here education. It was hypothesized that the results of this analysis might be quite different from the others. The chapter closes by comparing the results of these analyses to results obtained in a similar study of socioeconomic and educational dimensions of Florida counties.

D Analysis Based on Raw Data

As was explained earlier, the raw scores were first standardized so that the scale for each variable was the same. Whenever the phrase "raw scores" is used in this study, it is used to designate standardized raw scores. The d's were then computed according to the formula already

described and were then transformed into a matrix of distance similarity indices (DSI) with values approaching 1.0 indicating greatest similarity and those closest to 0.0 least similarity. The matrix was factor analyzed the same way an intercorrelation matrix would be, with the highest coefficients used in the diagonals as communality estimates. Provision was made for iteration but no iterations were necessary to achieve satisfactory convergence between the original communality estimates and the row sums of squared loadings. The resulting 27 principal axes account for 90 percent of the total score variance. Seven factors, accounting for 83 percent of the total score variance, 93 percent of the common variance, were rotated both orthogonally to the Varimax criterion and obliquely using the Simple Loadings procedure. As shown in Table 9, all but four of the 21 intercorrelations of the factors were above .30, and one third of them were over .40. The oblique solution was therefore accepted as the one most likely to approach simple structure. Rotations were also made according to the Maxplane and Oblimin criteria, but these proved inferior to the Simple Loadings solution. The complete primary solutions of these three rotation procedures will be shown in the Appendix. A summary of the Simple Loadings solution may be found in Table 10 and is the one which will be discussed below. In this and all subsequent rotated factor solutions, the factors are arranged in order of the size of their sums of squared loadings.

TABLE 9
Intercorrelations of Simple Loadings Primary Factors
Q Analysis Based on Raw Score Distances

Factor	1	2	3	4	5	6	7
1		.46	.31	.32	.33	.35	.31
2			.36	.58	.33	.37	.42
3				.41	.36	.60	.33
4					.56	.38	.23
5						.28	.12
6							.24
7							

Factor Labels

1 South
2 Plains
3 Eastern Seaboard
4 Far West
5 Young States
6 Midwest
7 Upper New England-Appalachia

The labels given these groups are indicative of "ideal" types rather than actual geographic regions. Consequently, a given state could group with Midwestern states and be part of the group "Midwest" without being actually located in this region. This practice of giving label names to designate the majority but not necessarily all of the cases within a given group is common. For example, when Russett (1967) designated a group of nation-states "Western Community," he included in it Japan, a nonwestern country.

In order to show the tie between the groups and the R analysis a table is presented with the discussion of each group giving the factor score levels of each state in that group on the five R factors. In discussing the Q factors

TABLE 10
Simple Loadings Primary Factors
Q Analysis Based on Raw Scores Distances

States	South 1	Plains 2	Eastern Seaboard 3	Far West 4	Young States 5	Midwest 6	Upper New England-Appalachia 7
South Carolina	1.00						
Alabama	.98						
Georgia	.92						
North Carolina	.89						
Mississippi	.88						
Arkansas	.81						
Tennessee	.76						
Louisiana	.70						
Virginia	.63						
Kentucky	.56						.38
Texas	.44				.31	.30	
South Dakota		.99					
North Dakota		.96					
Nebraska		.92					
Iowa		.84					
Montana		.68					
Kansas		.61					
Idaho		.59					
Wyoming		.51		.49			
Minnesota		.43		.32			
Oklahoma	.30	.33					
Massachusetts			.93				
Rhode Island			.87				
New Jersey			.83				
Connecticut			.80				

TABLE 10
Continued

States	South 1	Plains 2	Eastern Seaboard 3	Far West 4	Young States 5	Midwest 6	Upper New England-Appalachia 7
New York			.67				
Pennsylvania			.46			.39	
Maryland			.39	.32			
Delaware			.34				
Utah				.70			
Washington				.70			
Oregon				.66			
Colorado				.54			
California				.36		.36	
Alaska					.65		
Nevada					.62		
Arizona				.31	.52		
Florida					.51		
New Mexico				.38	.42		
Hawaii					.42		
Ohio			.30			.51	
Illinois						.49	
Indiana						.48	
Missouri						.42	
Michigan				.36		.41	
Wisconsin			.33			.33	
Maine							.58
New Hampshire			.48				.51
Vermont		.33					.49
West Virginia	.47						.48

the term "group" is used rather than "factor" in order to distinguish readily between R-type factors and Q-type factors.

Group 1, composed of the Southern states, is shown in Table 11. With more homogeneity than is true of the

TABLE 11
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 1: The South

State	Factor Loading	R FACTORS Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
South Carolina	1.00	--	---	0	0	-
Alabama	.98	-	--	0	-	0
Georgia	.92	-	--	0	-	-
North Carolina	.89	--	--	-	0	0
Mississippi	.88	--	---	0	-	++
Arkansas	.81	--	--	0	-	+
Tennessee	.76	--	-	-	-	0
Louisiana	.70	+	---	+	-	-
Virginia	.63	-	--	-	+	-
Kentucky	.56	--	0	--	0	+
West Virginia	.47	---	+	-	0	--
Texas	.44	+	-	0	-	+

FACTOR SCORE LEVELS

++ = +1.00 to +1.99	- = - .26 to - .99
+ = + .26 to + .99	-- = -1.00 to -1.99
0 = - .25 to + .25	--- = -2.00 to -2.99

other groups of states, these states have moderate to high negative scores on Factor 1, Socioeconomic, scores close to the mean on Educational Expenses vs. Manufacturing, and low negative scores on Frontier and, except for Mississippi, Agriculture. The most striking characteristic of this group, however, is the pattern of high negative scores on Factor 2, Racial-Educational. These states represent the high-percentage-

Black end of this bipolar factor, with extreme scores on all or most of the racial, educational, and cultural variables that make up Factor 2. It comes as no surprise that Mississippi has the most extreme score on this factor, representing a larger percentage of Black than any other state.

Group 2, shown on Table 12, is labeled the Plains and presents a profile which is quite different from that of

TABLE 12
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 2: The Plains

State	Factor Loading	R FACTORS Socio-Econ.	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
South Dakota	.99	-	+	++	0	+++
North Dakota	.96	-	+	++	0	++
Nebraska	.92	0	+	0	0	+++
Iowa	.84	-	++	0	-	+++
Montana	.68	-	+	++	+	+
Kansas	.61	0	+	+	-	+
Idaho	.59	--	++	+	0	+
Wyoming	.51	+	+	+++	0	0
Minnesota	.43	-	+	+	-	+
Oklahoma	.33	-	0	+	-	+

FACTOR SCORE LEVELS

+++ = +2.00 to +2.99 0 = - .25 to + .25
++ = +1.00 to +1.99 - = - .26 to - .99
+ = + .26 to + .99 -- = -1.00 to -1.99

Group 1. The states in this group are the only ones which have high positive scores on Agriculture. Most of these states have low negative scores on Socioeconomic. Only Idaho has a negative score on this factor greater than -1.0. The states in this group have moderate positive score on Racial-

Educational, indicating a largely White society, and similar scores on Educational Expenses vs. Manufacturing. Graphing these states on the variables which make up Educational Expenses Vs. Manufacturing reveals that their positive relation to this factor cannot be attributed to the positive end of this factor--high educational expenses--but rather to the negative end--low involvement in manufacturing. Table 9 shows that the correlation between this group and Group 4, the Far West, is .59, that is, approximately one third of the variance ($.59^2$) of the plains is shared by the states loading on Far West.

The factor analytic results of Group 3, the Eastern Seaboard, shown on Table 13, must represent a geographer's

TABLE 13
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 3: The Eastern Seaboard

State	Factor Loading	R FACTORS: Socio-Econ.	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
Massachusetts	.93	+	+	-	-	--
Rhode Island	.87	+	+	-	-	--
New Jersey	.83	++	0	-	-	-
Connecticut	.80	++	+	-	0	--
New York	.67	+++	-	-	-	0
Illinois	.48	++	0	--	0	++
New Hampshire	.48	--	++	--	0	--
Pennsylvania	.46	+	+	--	-	0
Maryland	.39	+	-	-	0	-
Delaware	.34	++	-	-	+	+

FACTOR SCORE LEVELS

+++ = +2.00 to +2.99
++ = +1.00 to +1.99
+ = + .26 to + .99
0 = - .25 to + .25
- = - .26 to - .99
-- = -1.00 to -1.99

dream. Omitting the states which load here but which have higher loadings in other groups (Illinois and New Hampshire), the states in this group arrange themselves in almost perfect latitudinal order, from north to south. Only New Jersey is slightly misplaced. These states tend to rank high on Socioeconomic, moderately high on Racial-Educational, moderately low on Educational Expenses vs. Manufacturing and Frontier (Delaware and Maryland are exceptions here, for they represent a population frontier), and from moderately low to very low on Agriculture. This group has a correlation of .60 with Group 6, the Midwest. Table 9 shows that these two groups are more highly correlated than any other two groups in this study.

One might be interested in knowing how the factor profiles of the states in this group differ from that of the urban state of California. As Table 10 shows, California does not clearly align with any one group, having low but sizeable (i.e., higher than .30) loadings on both Far West and Midwest. California has nearly as high a score on Socioeconomic as New York has, yet the rest of its profile differs. It has a score on Educational Expenses vs. Manufacturing near the mean and, in contrast to the Eastern Seaboard states, has a positive score on Agriculture. Its pattern is closest to that of Michigan, but even here there is no close alignment.

Group 4, the Far West, represents a small and loose cluster of states. It is shown on Table 14. Except for California's high score on Socioeconomic, these states have

TABLE 14
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 4: The Far West

State	Factor Loading	R FACTORS: Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
Utah	.70	0	+	++	-	--
Washington	.70	+	+	+	0	-
Oregon	.66	0	+	+	0	-
Colorado	.54	+	+	++	-	-
Wyoming	.49	+	+	+++	0	0
New Mexico	.38	+	-	++	0	-
California	.36	+++	0	0	-	+

FACTOR SCORE LEVELS
+++ = +2.00 to +2.99 0 = - .25 to + .25
++ = +1.00 to +1.99 - = - .26 to - .99
+ = + .26 to + .99 -- = -1.00 to -1.99

low positive scores on Socioeconomic and Racial-Educational. Some of their scores on Educational Expense vs. Manufacturing are higher than those of any other group and all have at least moderate positive scores. They cluster around the mean on Frontier and have widely ranging scores on Agriculture.

The states that are brought together in Group 5, shown on Table 15, do not represent a contiguous homogeneous group. They instead collectively make up perhaps America's last land "frontier." This includes Alaska, the Southwest, Florida, and Hawaii. As a group these states have moderate positive scores on Socioeconomic, moderate negative scores on Racial-Educational, widely ranging scores on Educational Expense vs. Manufacturing and Agriculture. Extensive graphing was done for these states on many of the 53 variables. Alaska

TABLE 15
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 5: Young States

		R FACTORS				
State	Factor Loading	Socio-Economic	Racial-Educ.	Educ. Expenses vs. Mfg.	Frontier	Agric.
Alaska	.65	0	-	-	++++++	-
Nevada	.62	+	0	+	++	-
Arizona	.52	+	-	++	0	-
Florida	.51	+	--	0	0	0
New Mexico	.42	+	-	++	0	-
Hawaii	.42	++	-	0	++	+

FACTOR SCORE LEVELS

++++++ = +5.00 to +5.99   0 = - .25 to + .25
++ = +1.00 to +1.99   - = - .26 to - .99
+ = + .26 to + .99   -- = -1.00 to -1.99

consistently was the state with extreme scores on many of the variables and is the pivot state for this group, that is, the state with the highest loading on this group. Alaska has a factor score of 5.98 on the R factor Frontier. 5.98 standard deviations above the mean.

Group 6 is made up of the Midwestern states. This group is shown on Table 16. Together these states have moderate positive scores on Socioeconomic and Racial-Educational, but their real characteristics are the high negative scores on Educational Expenses vs. Manufacturing. The "negative" scores on this factor result both from low educational expenses per capita in these states and proportionally high involvement in manufacturing. Group 6 has a greater emphasis upon manufacturing and a smaller percentage of the economy spent

TABLE 16
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Score Distances
Group 6: The Midwest

State	Factor Loading	Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
		R FACTORS				
Ohio	.51	-	+	--	-	+
Illinois	.49	++	0	--	0	++
Indiana	.48	0	+	--	0	+
Missouri	.42	0	+	--	-	++
Michigan	.41	+	0	-	-	0
Pennsylvania	.39	+	+	--	-	0
California	.36	+++	0	0	-	+
Wisconsin	.33	+	+	--	0	+

FACTOR SCORE LEVELS

+++ = +2.00 to +2.99
++ = +1.00 to +1.99
+ = + .26 to + .99
0 = - .25 to + .25
- = - .26 to - .99
-- = -1.00 to -1.99

on education than the other groups of states have. The Midwestern states have medium range scores on Frontier and, except for Pennsylvania, medium to high positive scores on Agriculture.

Group 7, identified as "Upper New England-Appalachia," is shown on Table 17. The point was made in the discussion of the R analysis that the states which contribute most of the variance to Factor 2, Racial-Education, rank near the bottom of the states when ranked according to their factor scores on Socioeconomic. Group 7 is made up of the states which rival the South in high negative scores on Socioeconomic but contrast with the South on scores on Racial-Educational. Group 7 thus represents the "White" end of this factor. Even though Kentucky, the lowest loading state

TABLE 17
Oblique Factor Loadings and R Factor Levels
Q Analysis Based on Raw Scores Distances
Group 7: Upper New England-Appalachia

State	Factor Loading	R FACTORS: Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
Maine	.58	--	++	-	0	--
New Hampshire	.51	--	++	--	0	--
Vermont	.49	--	++	0	0	-
West Virginia	.48	---	+	-	0	--
Kentucky	.38	--	0	--	0	+

on this group, has a score on Racial-Educational that is near the mean, its profile is very similar to that of the other states in this group. The states in Group 7 have moderate negative scores on Educational Expenses vs. Manufacturing, medium range scores on Frontier, and negative scores on Agriculture.

As has been mentioned, ofttimes when researchers make use of Q analysis they used as data, not the raw scores, but factor scores obtained from an R-type analysis. Rummel (1970) argues that the use of raw data artificially weights correlated variables. The use of factor scores for data in a Q analysis, however, presupposes that each factor is of equal importance. This might be a justifiable assumption if one were dealing with such general factors as Social, Political, Economic, or Educational. In this study, however, the R factors are not of equal importance. Factor 1, Socio-economic, is a general factor that includes variables from

many realms. Can one really say that Frontier, accounting for 12 percent of the total score variance, or Agriculture, accounting for 9 percent, is equal in importance to Socio-economic, accounting for 22 percent? This assumption would have to be made if Rummel's argument for the use of factor scores is valid.

One way to distinguish between the factors is to weight each one according to some criterion. Empirically, the only criterion which presents itself is that of weighting each factor according to the percentage of variance each explains. In some research an external criterion might be devised to test the relative importance of the factors. Below are presented for comparative purposes, first, the results of a Q analysis based on d's from unit weighted factor scores, second, on d's from factor scores weighted according to the percentage of variance explained by each factor, and, third, on d's from factor scores weighted according to the percentage of educational variables in each factor.

D Analysis Based on Unit Weighted Factor Scores

Using factor scores as the data for D analysis did not prove as simple as might be supposed, mainly because getting the factor scores themselves proved difficult. As stated in Chapter 2, the inverse of the correlation matrix is a necessary step in obtaining factor score estimates according to the complete method. The computation of the inverse proved impossible for these data, with the off-

diagonal values in the resulting matrix much too large. To get around the problem a return was made to the fundamental equation in factor analysis: an orthogonal factor matrix multiplied by its transpose reproduces the correlation matrix, FF' = R. The eight factor Varimax solution for the E analysis, which was analyzed in Chapter 3, was postmultiplied by its transpose. The values in the resulting matrix were very close to those in the original correlation matrix, with the difference that the inverse, and subsequently the factor score estimates, could be computed.

Using these factor scores as replicates, the distance analysis program was run. Standardization was not necessary since the factor scores were already standardized to a mean of 0.0 and a standard deviation of 1.0. Twenty-seven principal axes, accounting for 93 percent of the total score variance, emerged. Seven of these factors, accounting for 89 percent of the total score variance and 95 percent of the common variance, were rotated both orthogonally and obliquely. As before, the Simple Loadings solution proved superior. A summary table of this solution appears on Table 18, and a complete table in the Appendix. On these and subsequent D analyses tables the groups are numbered rather than labeled. Cross-comparison with D analysis using raw scores can be made by looking at the states that are included in each group. This was done because the former labels did not always fit the new groups.

Relevant here are the differences between this grouping and the one based on raw scores. New York is now the pivot

TABLE 18
Simple Loadings Factors
Q Analysis Based on Unit Weighted Factor Score Distances

States	1	2	3	4	5	6	7
South Carolina	.94						
Alabama	.93						
Georgia	.90						
Mississippi	.84						
North Carolina	.79						
Arkansas	.76						
Pennsylvania	.69						
Louisiana	.67	.33					
Virginia	.65						
Texas	.50						
Florida	.50						
Kentucky	.46		.32			.37	
Utah		.79					
Wyoming		.78					
Colorado		.75					
New Mexico		.75					
Oregon		.69					
Arizona		.64					
Washington		.62					
Montana		.50		.44			
Oklahoma	.30	.39					
Indiana			.83				
Missouri			.82				
Wisconsin			.81				
Ohio			.75				
Pennsylvania			.65				
Illinois			.60		.37		
Minnesota		.36	.37	.36			
South Dakota				.96			
North Dakota				.94			
Nebraska				.84			
Iowa				.78			
Idaho		.36		.46			
Wyoming		.40		.46			
New York					.75		
New Jersey					.73		
California					.60		
Connecticut					.60	.30	
Massachusetts					.53	.51	
Delaware					.47		.34
Maryland					.46		
Michigan			.42		.42		

TABLE 18
Continued

State	1	2	3	4	5	6	7
Maine						.81	
New Hampshire						.78	
West Virginia						.76	
Vermont						.65	
Rhode Island					.49	.59	
Alaska							.66
Nevada		.31					.59
Hawaii							.55

state for the group of states formerly labeled Eastern Seaboard. This puts California, which has as high a score on Socio-economic as New York has, in this group. California's scores are now compared with New York rather than with Massachusetts. Arizona and New Mexico now group with the Far Western states rather than with Alaska, Nevada, and Hawaii. This shift illustrates the change from raw scores to equally-weighted factor scores. As shown by Table 19, New Mexico, Arizona,

TABLE 19
E Factor Scores of Selected States

State	Socio-Econ.	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
Utah	.17	.77	1.93	-.83	-1.46
New Mexico	.30	-.68	1.90	.15	-.92
Arizona	.86	-.69	1.49	.10	-.35
Alaska	-.06	-.28	-.49	5.98	-.84

and Alaska have scores on Factor 2 that are comparatively similar, whereas New Mexico, Arizona, and Utah have equally high scores on Factor 3. The weights of these two factors, however, are different in the two analyses. In the raw score analysis how states pattern on variables which make up Factor 2 is more important than how they pattern on Factor 3 variables, and, consequently, New Mexico and Arizona group with Alaska. In the unit weighted factor score analysis both factors are equal in importance, and Arizona and New Mexico group with the rest of the Southwest. For a similar reason Florida now groups with the South. Nevada, however, continues to group with Alaska and Hawaii, just as Florida did in the raw score analysis.

Although this comparison illustrates the difference that results from the use of the two types of data, it also oversimplifies the difference. The raw score analysis retains the variance of each state on each variable, and these variance differences alone, even without a shift in weighting given the factors, can cause a state to change from one group to another. This will be seen in the discussion of D analysis based on factor scores weighted according to variance, where the proportionate share of variance accounted for by each factor is retained, yet still resulting in some shifting of states from one group to another.

Over all, however, the differences in the groupings of the two analyses are not great. One would not be justified on the basis of these data, in condemning the use of unit

weighted factor scores despite reservations about the rationale upon which the method is based.

D Analysis Based on Factor Scores Weighted on Variance

In order to prepare the data for this analysis the matrix of factor scores was multiplied by the vector of percentages of common factor variance accounted for by each factor. This meant that Factors 1 and 2, accounting for 30 and 26 percent, respectively, of the variance explained by the five interpretable R factors, played dominant roles in this analysis. The distance analysis program was then run, but without the usual option of standardizing scores, since this would in effect offset the use of weights. The capsule seven factor oblique solution, accounting for 86 percent of the total score variance, is shown on Table 20; the completed solution is in the Appendix.

It was hypothesized that this analysis would be more like the results of the one based on standardized raw scores than on the one on unit weighted factor scores because the original factor variance was retained. A comparison of Tables 10, 18, and 20 shows that this did not prove true. The two employing factor scores are closer to each other than either is to the one using raw scores, and none of the three is greatly different from the others.

A few differences should be noted. With the greatest weight given R Factor 1, California and Illinois now both

TABLE 20
Simple Loadings Factors
Q Analysis Based on Variance Weighted Factor Score Distances

State	1	2	3	4	5	6	7
Alabama	.96						
South Carolina	.96						
Mississippi	.95						
Georgia	.93						
North Carolina	.80						
Arkansas	.78						
Tennessee	.71						
Virginia	.67						
Louisiana	.67						
Texas	.49	.33					
Florida	.47	.31					
Oklahoma	.33		.30	.30			
New York		1.00					
California		.94					
New Jersey		.92					
Delaware		.77					
Illinois		.75					
Connecticut		.55				.32	
Maryland		.53					
Michigan		.52				.37	
Utah			.77				
Colorado			.68				
Wyoming			.67	.32			
Oregon			.63				
Washington			.62				
New Mexico	.31		.59				
Massachusetts			.50			.44	
Arizona		.32	.49				
Minnesota			.38	.36		.35	
South Dakota				.83			
North Dakota				.83			
Nebraska				.74			
Iowa				.71		.33	
Montana			.38	.55			
Idaho				.53	.45		
Wyoming			.37	.52			
Maine					.94		
West Virginia					.85		
Vermont					.81		
New Hampshire					.78		
Kentucky	.41				.58		

TABLE 20
Continued

State	1	2	3	4	5	6	7
Pennsylvania						.72	
Indiana						.71	
Ohio						.70	
Wisconsin						.69	
Missouri						.65	
Rhode Island			.42			.49	
Alaska							.69
Nevada							.57
Hawaii		.46					.54

group with the urban East. Rhode Island has shifted in each of the analyses. It first grouped with the Eastern Seaboard, next was split between the groups formerly labeled Upper New England-Appalachia and the Eastern Seaboard, and in this present analysis is split, along with Massachusetts and Minnesota, between the Midwest and the Far West.

D Analysis Based on Factor Scores Weighted on Percentage of Educational Variables

This analysis is presented as an example of how factor scores might be weighted to take into account the influence of a certain aspect of socioculture, here education. First the variables were divided into educational or school related variables and noneducational variables. Those classified as educational variables are so indicated in the list of variables in the Appendix. This division was an arbitrary one both because many of the variables were complex and

because the choice was not made on the basis of theory. For instance, is the variable, average teacher salary, an educational variable or an economic one? This particular variable was included among the educational variables even though it may likely be that teacher salaries are more an indication of local cost of living, or even the law of supply and demand (witness the relatively low teacher salaries in many college communities where the supply includes student and faculty wives), than an indication of community effort or lack of effort to provide quality education.

After the variables had been so divided, factor weights were determined according to the following equation:

wgt = ed/total

wgt - weight for that factor

ed - total of the squared loadings of educational variables with loadings greater than .30 on that factor

total - total of the squared loadings of all variables with loadings greater than .30 on that factor.

The resulting weights were as follows: Socioeconomic, .42; Racial-Educational, .55; Educational Expenses vs. Manufacturing, .47; Frontier, .39; and Agriculture, .28. Proportionally, the end results of all these calculations were weights not greatly different from those based on percentage of variance accounted for. This is unfortunate since the emphasis of this fourth D analysis was to be on showing grouping differences between this and the former analyses,

but the similarity could not have been predetermined. This points out how all pervasive education is. No matter what the factor label was, educational variables were an important part of each of the five factors.

This vector of weights was multiplied times the matrix of factor scores and the D analysis program was run as before. A summary of an oblique rotation of seven factors, accounting for 88 percent of the total score variance, may be found in Table 21. The groupings are similar to those cited before. Illinois now loads on the Midwest as well as on the urban states, whereas before it grouped exclusively with the urban states. Minnesota, formerly with variance split between three groups, now groups only with the Plains, though with only a moderate loading. Massachusetts and Rhode Island split their variance between the New York-type states and the Maine-type states. All in all, however, the groupings on factor scores weighted by percentages of educational variables are not greatly different from those which were introduced earlier.

One way an overall comparison can be made of these groupings is to look at the pivot states of the four analyses, as shown by Table 22. The labels and order of the groups are taken from the D analysis based on raw scores. As Table 22 shows, the pivot states stay essentially the same, with only slight shifts. With the pivot states the same or similar, the groupings of states themselves are similar. The group of states with large DSI from South Carolina, for instance,

TABLE 21
Simple Loadings Factors
Distance Analysis Based on Factor Scores
Weighted According to Percentage of Educational Variables

State	1	2	3	4	5	6	7
South Carolina	.95						
Alabama	.95						
Georgia	.92						
Mississippi	.91						
Arkansas	.80						
North Carolina	.80						
Louisiana	.70				.32		
Tennessee	.70						
Virginia	.67						
Texas	.53			.33			
Florida	.52			.31			
Kentucky	.43	.33				.38	
Indiana		.93					
Missouri		.92					
Ohio		.90					
Wisconsin		.87					
Pennsylvania		.86					
Illinois		.56		.44			
South Dakota			.86				
North Dakota			.83				
Nebraska			.79				
Iowa			.75				
Idaho			.58			.37	
Wyoming			.58				
Montana			.54		.39		
Minnesota			.47				
Oklahoma	.30		.31				
New York				.89			
New Jersey				.82			
California				.75			
Delaware				.59			.32
Connecticut				.58			
Maryland				.52			
Michigan		.38		.50			
Massachusetts				.46		.45	
New Mexico					.77		
Utah					.74		
Wyoming					.73		
Arizona					.65		
Colorado					.61		
Oregon					.51		
Washington					.46		

TABLE 21
Continued

State	1	2	3	4	5	6	7
Maine						.86	
New Hampshire						.77	
West Virginia						.76	
Vermont						.76	
Rhode Island				.37		.50	
Alaska							.65
Nevada					.35		.56
Hawaii				.33			.54

TABLE 22
Pivot States for Four Distance Analyses

Factors	Raw Score	Unit Weights	Variance	Education
South	S. Caro.	S. Caro.	Alabama	S. Caro.
Plains	S. Dakota	S. Dakota	S. Dakota	S. Dakota
Eastern Seaboard	Mass.	N. York	N. York	N. York
Far West	Utah	Utah	Utah	N. Mexico
Young States	Alaska	Alaska	Alaska	Alaska
Upper New Eng.-Appalachia	Maine	Maine	Maine	Maine

is very nearly the same for all four of the analyses. Two reasons may be given for the similarity in groupings. One is that the factor structure itself was stable enough to produce similar groupings no matter, within reason, how the factor scores were weighted. In other words, New Jersey would be like New York and Alabama like South Carolina, no matter how the states were looked at. This would speak for a regional approach to grouping since with only a few exceptions these groups are regional.

The other reason has already been mentionedin comparing variance weights with educational weights: none of the weightings used made much real difference in the data. That is to say, factor scores weighted on percentage of variance explained by each factor and percentage of educational variables loading on each factor were not appreciably different from either each other or from the other two types of data used. This can best be seen in Table 23, a direct comparison of the proportional weights of the three analyses employing factor scores.

TABLE 23
Proportional Weights Used in Three Factor Score Analyses

Factor	Unit Weights	Variance Weights	Educational Weights
Socioeconomic	.20	.29	.20
Racial-Educational	.20	.25	.26
Educ. Expenses Vs. Mfg.	.20	.18	.22
Frontier	.20	.16	.18
Agriculture	.20	.11	.13

It would seem from these analyses that the use of factor scores in D analyses, whether given equal weight or weighted on variance explained, results in groupings similar to those obtained from the use of standardized raw scores. One might conclude from these analyses that the use of factor scores is advisable because the method is both simpler and cheaper. However, other research examples may have quite different results. Below is a summary of the factor analytic study of characteristics of Florida counties referred to earlier.

D Analyses Based on Florida Data

A study similar to that of the 50 states was made and reported of the grouping of the 67 Florida counties on 74 socioeconomic, political, and educational variables (Chang, 1971). An R-type analysis yielded 15 interpretable factors. These factors and the percentage of total score variance accounted for by each are presented in Table 24. The factors Socioeconomic and Agriculture correspond to their counterparts in the United States study. Socioeconomic explains 25.9 percent of the total score variance. The remaining variance is divided among 14 additional factors, whereas remaining variance is divided among only four additional factors in the present study. This difference goes a long way toward explaining other differences in the results of the two studies which will be considered below.

Three distance analyses were made, one using as data the standardized raw scores, the second unit weighted factor

TABLE 24
Socioeconomic, Political, and Educational Dimensions of Florida Counties

Factor	% of Total Score Variance
Socioeconomic	25.9
Expenditure Per Pupil	7.5
Nonwhite	5.4
Agriculture	5.3
Retired	5.3
Community Support for Schools	4.2
Enrolled in High School	4.2
Population Growth	3.4
Manufacturing	3.4
University Centers	3.3
Little Change in Per Capita Income	3.0
Increase in School Revenue	2.5
Tourism	2.5
Teachers on Continuing Contract	2.4
Millage Required for School Revenue	2.4
	80.7

scores, and the third variance weighted factor scores.

A summary of the oblique rotation of the analyses employing raw scores is shown in Table 25. The table shows seven factors, accounting for 81 percent of the total score variance. In terms of the R factors, counties with sizeable loadings on Group 1 tend to have high scores on Socioeconomic, and the counties with the highest loadings, have high scores also on such factors as Retired, Population Growth, Community Support for Schools, and Tourism. These are the tourist counties, the boom counties. Group 2 represents counties at the other extreme, that is, counties which have very low scores on socioeconomic characteristics. These have

TABLE 25
Simple Loadings Factors
Q Analysis Based on Raw Score Dist[illegible]
Florida Counties

[illegible]	1	2	3	4	5	6	[illegible]
Sarasota	.94						
Palm Beach	.84						
Pinellas	.80						
Charlotte	.80						
Volusia	.72						
Broward	.70						
Dade	.66						
Manatee	.60						
Lake	.59		.31				
Indian River	.57						
Collier	.54						.31
Lee	.53		.32				
Orange	.50		.31				
Hillsborough	.48			.35			
Highlands	.48		.35				
Martin	.47		.33				.31
St. Lucie	.44		.36				
Osceola	.42		.39				
St. Johns	.40						
Hernando	.40						.35
Duval	.38			.34			
Monroe	.36		.34				
Marion	(.29)					(.29)	
Holmes		.99					
Washington		.80					
Walton		.76					
Lafayette		.73					
Calhoun		.69					
Gilchrist		.68					
Jackson		.59				.34	
Liberty		.56					
Suwannee		.52					
Baker		.50	.30				
Wakulla		.46					.33
Sumter		.46					
Union		.40					
Levy		.39					
Bradford		.38					
Franklin		.33					
Hardee			.76				
Okeechobee			.68				
Hendry			.58				
Clay			.55				

TABLE 25
Continued

County	1	2	3	4	5	6	7
Okaloosa			.55		.37		
De Soto			.52				
Pasco	.38	.31	.47				
Seminole			.44				
Polk	.34		.38				
Taylor				.70			
Gulf				.69			
Nassau				.55			
Putnam				.52			
Dixie		.33		.52			.31
Santa Rosa				.47			
Escambia				.38			
Bay				.35			
Brevard				.33			
Leon					.99		
Alachua					.95		
Gadsden						.70	
Jefferson						.66	
Madison		.41				.54	
Columbia		.31				.37	
Flagler						.30	
Glades							.47
Citrus	.45						.45
Hamilton						.32	.39

the northern counties which have few tourists, few retirees, relatively smaller population and economic growth. Group 2 may be contrasted with Group 6, also in the northern tier, by the former's lower socioeconomic level and the latter's larger percentage of nonwhite population. In this study, the state's poorest counties represent a distinct grouping from those counties which have the highest population of nonwhite. Counties in Group 3 are those which depend upon

agriculture to a larger degree than do the other counties. By contrast, Group 4 includes counties which depend more upon manufacturing and such less upon agriculture. Group 5 consists of Alachua and Leon. These two counties, where the University of Florida and Florida States University, respectively, are located, have very high scores on the factor University Centers, and scores on all the other factors which are close to each other.

The distance analysis based on unit weighted factor scores for these Florida data illustrates well the potential differences in results between using factor scores as data and using raw scores. As has been mentioned, in contrast to the five interpretable R factors in the United States study, the Florida study yielded 15. Giving equal weight to each factor results in Socioeconomic being given, along with all other factors, a weight of 6.7 percent instead of the original 25.9. The 16 factor oblique solution, accounting for 88 percent of the total score variance, is presented in Table 26. By referring to the factor scores of the counties within the various groups one can understand the relationship between the factors and the R factor scores which served as data. Several factor descriptions may be cited as illustrations. Group 1 counties are in large part those which made up Group 6 of the D analysis based on raw scores, which are characterized by moderately low scores on Socioeconomic and high scores on Nonwhite. Group 2 is made up of Alachua and Leon, the two university centers. Group 3 includes counties which

TABLE 26
Simple Loadings Factors
Q Analysis Based on Unit Weighted Factor Score Distances
Florida Counties

County	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gadsden	1.00															
Jefferson	.88															
Madison	.86															
Flagler	.52															
Jackson	.56															
St. Lucie	.45		.32													
Columbia	.35															
Bradford	.27															
Alachua		1.04														
Leon		1.02														
Osceola			.70													
Highlands			.52					.32								
Clay			.43													
DeSoto			.43													
Martin			.42						.34							.34
Indian River			.41													
Hendry			.37					.34								
Seminole			.34													
Taylor				.68												
Gulf				.67												
Liberty				.64												
Putnam				.50												
Dixie				.44												
Calhoun				.40				.34								
Santa Rosa				.37			.32									

TABLE 26
Continued

County	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Dade					.95											
Duval					.68											
Hillsborough					.58											
Escambia					.40								-.33			
Palm Beach					.34					.34						
Polk					.31											
Nassau						.68										
St. Johns						.68										
Volusia						.62										
Bay						.42										
Marion						.34										
Lee						.28										
Brevard							.85									
Okaloosa							.66									
Orange							.34									
Lafayette								.78								
Gilchrist								.59								
Manatee									.71							
Pasco									.54							
Holmes									.47				.42			
Sarasota									.43							
Washington									.30							
Collier										.80			.47			
Hardee										.51	.33	-.32				
Monroe						.33				.37						

TABLE 26
Continued

County	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Wakulla							.33				.64					
Union											.37				-.34	
Levy											.31					
Sumter											.24					
Hamilton												.62				
Suwannee												.50	.40			
Citrus											.33	.42				
Hernando												.29				
Charlotte										.31			-.36			
Baker													.34			
Walton													.34			
Okeechobee													.29			
Pinellas														.69		
Broward							.40							.35	.48	
Franklin			.31								.32				.48	
Glades																.73

have a high socioeconomic level, but have lower scores on Population Growth and Retired. Counties in Group 5 have high scores on the factor Tourism. Other groups could be characterized similarly in terms of factor scores.

Here, unlike the results of the United States study, equating the R factors produces greatly different results than were obtained by the used of raw scores. By making the assumption, for instance, that Socioeconomic is no more important than Increase in School Revenue or Teachers on Continuing Contract, some rather strange groupings of counties emerged. It is difficult to conceive how the groupings of counties shown in Table 26 could be used for any purpose.

These groupings may be contrasted with the ones resulting from the use of factor scores based on percentages of variance explained. As before, the hypothesis was that this type of distance analysis would most closely resemble the one based on raw scores. This seems to be true, as seen in Table 27, which presents the six factor oblique solution, accounting for 88 percent of the total score variance. But while the groupings have resemblances to the one presented in Table 25, there are some noticeable differences. Factor 1, Socioeconomic, accounts for 30 percent of the common factor variance of the R analysis; therefore, it has preponderant influence upon the factor structure presented in Table 27. This may be seen by comparing groupings with the factor score for each county on Factor 1, given in the last column. Group 1 is made up of counties with medium range Socioeconomic

TABLE 27
Simple Loadings Factors
Q Analysis Based on Variance Weighted Factor Score Distances
and Scores on Factor 1
Florida Counties

County	1	2	3	4	5	6	Factor Score Socioeconomic
Lee	.95						.59
Polk	.92						.71
Martin	.92						.43
Seminole	.91						.38
Highlands	.86						.35
Leon	.85						.60
Indian River	.84						.70
Volusia	.84						.69
Lake	.82						.70
Alachua	.82						.55
St. Lucie	.81						.41
Hendry	.80						.77
Bay	.79						.40
Escambia	.77						1.11
Monroe	.79						.53
St. Johns	.75						.22
Manatee	.71						.35
Okaloosa	.70		.32				1.13
Osceola	.67				.31		.06
Okeechobee	.61						-.03
Hillsborough	.60		.44				1.18
Clay	.56				.36		-.10
DeSoto	.53				.36		-.13
Gulf	.49						-.02
Nassau	.44				.33		-.13
Hardee	.44						-.30
Collier	.40					.34	.22
Putnam	.34			.34			-.17
Pasco	.34				.32		-.35
Holmes		.99					-1.69
Calhoun		.99					-1.58
Dixie		.94					-1.55
Wakulla		.91					-1.71
Gilchrist		.88					-1.19
Washington		.87					-1.19
Liberty		.84					-1.23
Union		.76					-1.15
Levy		.75					-1.02
Walton		.71					-1.02
Citrus		.60					-.85
Hamilton		.58		.38			-1.22
Lafayette		.51					-.65

TABLE 27
Continued

County	1	2	3	4	5	6	Factor Score Socioeconomic
Suwannee		.45		.30			-.83
Sumter		.43					-.74
Santa Rosa	.32	.36			.34		-.32
Baker		.36		.35	.32		-.72
Taylor		.34					-.47
Franklin		.33			.32		-.50
Brevard			.96				2.59
Broward			.91				2.02
Dade			.90				2.65
Orange			.79				1.77
Pinellas			.73				1.54
Duval	.30		.73				1.71
Palm Beach	.33		.53				1.24
Sarasota	.43		.53				1.24
Gadsden				.93			-.67
Jefferson				.84			-.89
Madison				.78			-.99
Jackson		.36		.51			-.78
Flagler				.50			-.49
Columbia				.49			-.43
Bradford				.40			-.55
Marion				.36	.31		-.31
Hernando					.29		-.49
Charlotte	.34					.41	.11
Glades						.36	-.22

factor scores, ranging from low positive to low negative. Group 2 counties range from low negative scores to moderate negative scores. Group 3 counties have high positive scores. Because the R factors consist of one predominant factor and a number of smaller factors, this one factor has overriding influence upon this analysis. To a degree, this is true in the D analysis based on raw scores, but that analysis at least retains the variance of the original variables.

use of factor scores ignores this variance and with it valuable information for grouping the cases on the variables.

These comparisons of D analyses employing different kinds of data may lead one to two quite different conclusions. As seen in the Florida data, the choice of data can make a tremendous difference in the grouping results. Use of unit weighted factor scores can lead to gross distortion if there is a sizeable difference between the variance explained by the individual factors and/or their substantive importance. Factor 5, Retired, for instance, cannot be compared in scope to Factor 1, Socioeconomic, because the former is only an isolated aspect of the human environment which happens not to correlate significantly with the general human environment, of which the latter is a chief measure. Factor 1 is general; Factor 5 is specific. Only with all specific factors or all general ones would one be justified in weighting each one equally. Even then, one would have to be willing to make a theoretical judgment that two factors were indeed equal in importance in order to give them equal weighting.

The states data shows, on the other hand, that the use of unit weighted factor scores cannot be condemned just because it sometimes leads to distortion. The groupings resulting from the use of equally weighted factor scores in this study were not greatly different from those produced by the use of raw data. If one is willing to equate the k factors, the use of unit weighted factor scores is justifiable.

in this case. And in studies where the R factors are not greatly different in percentage of variance accounted for, there may be little difference between the results from the use of raw scores and the use of raw factor scores.

In most research, however, it would seem that if one is to use factor scores, some weighting criterion other than unit weights should be used, probably most often percentage of variance accounted for. This most nearly reproduces the results obtained by the use of raw scores. The use of weights based on the percentage of educational variables is offered as an example of an alternate weighting criterion, one that could be used in developing or testing theory. In focusing on a specific problem, one might use an even more complicated weighting equation, such as giving educational support variables on the factors greater weights and other educational variables lesser weights.

Unless one is focusing on such a specific domain, however, the real question is, why use factor scores at all? There might be two possible justifications, one cost and the other ease of interpretation. One might choose to use factor scores if one is using a program in which the use of a large number of variables is prohibitly costly in comparison to using a small number of factor scores. This was not true in this study. The use of a 50 by 53 matrix, though more expensive than the use of a 50 by 5 matrix, was not greatly so. Secondly, it must be admitted that it is easier to explain groups in terms of five factors than in terms of 53

variables, as in the Florida analysis, 74. One needs a method beyond distance analysis to somehow tie the groups to the actual variable scores which explain the difference between groups of cases. One technique for doing this is presented in the next chapter.

The biggest criticism for the use of factor scores, no matter what weighting is used, is the loss of information that results. Variance on individual variables is valuable in grouping cases. As has been pointed out before, cases with similar factor scores still may have differences in levels of individual variables which can be capitalized on in grouping. One is never justified in using factor scores indiscriminately. Only when the alternatives are carefully weighed and the limitations of each one considered, should one use factor scores in preference to raw scores.

CHAPTER 5

Q ANALYSIS BASED ON CORRELATION

In the transpose analysis the scores for each characteristic are first standardized in order to eliminate variable scale differences. Each case's scores are then intercorrelated with those of each other case, resulting in a cases by cases matrix, in this instance, a 50 by 50 matrix, with the information on individual variables or characteristics being lost. Highest correlations serve as original communality estimates. The intercorrelation matrix is then factor analyzed and rotated to produce factors, which are in reality groups of cases which vary together. In this study the eight factor Varimax solution, accounting for 82 percent of the total score variance, 86 percent of the common variance, proved most stable. The factors were also rotated obliquely, but they proved essentially independent. The complete Varimax solution is shown in the Appendix.

Group 1, presented in Table 28, is a bipolar factor with the Southern states at one end and certain Western states at the other. The states at the two poles of this group vary especially on the socioeconomic, racial, and educational characteristics that make up R Factors 1, 2, and 3. The negative or, so to speak, Western end of this pole does not include those states which ranked at the opposite end of Factor 2, Social-Educational, from the South

TABLE 28
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation
Group 1

State	Factor Loading	Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
		R FACTORS				
Arkansas	.89	--	--	0	-	+
Alabama	.88	-	--	0	-	-
South Carolina	.88	--	---	0	0	-
Tennessee	.87	--	-	-	-	+
North Carolina	.86	--	--	-	0	0
Mississippi	.85	--	---	0	-	++
Georgia	.81	-	--	0	-	-
Kentucky	.79	--	0	--	0	+
Virginia	.65	-	--	-	+	-
West Virginia	.52	---	+	-	0	--
Wyoming	-.46	+	+	+++	0	0
Minnesota	-.53	+	+	+	-	+
Connecticut	-.56	++	+	-	0	--
California	-.62	+++	0	0	-	+
Colorado	-.71	+	+	++	-	-
Oregon	-.75	0	+	+	0	-
Washington	-.83	+	+	+	0	-

FACTOR SCORE LEVELS

+++ = +2.00 to +2.99	0 = - .25 to + .25
++ = +1.00 to +1.99	- = - .26 to - .99
+ = + .26 to + .99	-- = -1.00 to -1.99
	--- = -2.00 to -2.99

because those states shared the South's low scores on the urban and socioeconomic variables. This analysis, based as it is on correlation, brings together polar states whose scores on the variables of these two important factors vary in opposite directions. Consequently, the states loading negatively on the South are those which tend to have high positive score on both Socioeconomic and Racial-Educational in contrast to the South's negative scores on these two

factors. In addition, these Western states have much higher scores on Educational Expenses vs. Manufacturing than do the Southern states.

Table 29 shows the states which make up Group 2. The states which load positively on this group tend to have

TABLE 29
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation
Group 2

		R FACTORS				
State	Factor Loading	Socio-Economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
Illinois	.75	++	0	--	0	++
Pennsylvania	.74	+	+	--	-	0
Missouri	.73	0	+	--	-	++
Rhode Island	.73	+	+	-	-	--
Massachusetts	.70	+	+	-	-	--
New York	.66	+++	-	-	-	0
Ohio	.62	+	+	--	-	+
New Jersey	.62	++	0	-	-	-
Wisconsin	.56	+	+	--	0	+
New Hampshire	.50	--	++	--	0	--
Alaska	-.53	0	-	-	+++++	-
Wyoming	-.56	+	+	+++	0	0
Arizona	-.67	+	-	++	0	-
Utah	-.71	0	+	++	-	--
New Mexico	-.79	+	-	++	0	-

FACTOR SCORE LEVELS

+++++ = +5.00 to +5.99   0 = - .25 to + .25
+++ = +2.00 to +2.99   - = - .26 to - .99
++ = +1.00 to +1.99   -- = -1.00 to -1.99
+ = + .26 to + .99

high positive scores on Factor 1, Socioeconomic, and Factor 2, Racial-Educational. What marks the difference between them and the states at the negative poles of both Group 1 and Group 2 are their low scores on Factor 3, Educational Expenses

vs. Manufacturing. The high socioeconomic states in Group 1 have widely ranging but positively leaning scores on Factor 3; the states at the negative pole of Group 2, except Alaska, have high positive scores on Factor 3.

Group 3, shown on Table 30, is made up of the Plains states. They have a profile of negative to medium scores

TABLE 30
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation
Group 3

		R FACTORS				
State	Factor Loading	Socio-economic	Racial-Educ.	Educ. Expense vs. Mfg.	Frontier	Agric.
South Dakota	.92	-	+	++	0	+++
Nebraska	.91	0	+	0	0	+++
Iowa	.87	-	+	0	0	+++
North Dakota	.86	-	++	0	-	+++
Montana	.74	-	+	++	0	++
Kansas	.72	0	+	++	+	+
Idaho	.65	--	+	+	-	+
Minnesota	.54	+	++	+	0	+
Minnesota			+	+	-	+

FACTOR SCORE LEVELS

+++ = +2.00 to +2.99   0 = - .25 to + .25
++ = +1.00 to +1.99   - = - .26 to - .99
+ = + .26 to + .99   -- = -1.00 to -1.99

on Socioeconomic, positive scores on Racial-Educational, positive scores frequently on Educational Expenses vs. Manufacturing, and, as the most distinguishing feature, high positive scores on Agriculture. This, like the South, is as easily distinguishable in Q-based-on-correlation as in Q-based-on-distance.

Group 4, Table 31, is comparable to Group 7 of the D analysis, Upper New England-Appalachia. As has been mentioned

TABLE 31
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation
Group 4

State	Factor Loading	R FACTORS: Socio-Econ.	Racial Educ.	Educ. Expense vs. Mf	Frontier	Agric.
Maine	.87	--	++	-	0	--
Vermont	.74	--	++	0	0	-
New Hampshire	.72	--	++	--	0	--
West Virginia	.48	---	+	-	0	--
Texas	-.54	+	-	0	-	+

FACTOR SCORE LEVELS

++ = +1.00 to +1.99
+ = + .26 to + .99
0 = - .25 to + .25
- = - .26 to - .99
-- = -1.00 to -1.99
--- = -2.00 to -2.99

several times, these states are characterized by negative scores on Socioeconomic and Agriculture and positive scores on Racial-Educational. Texas loads negatively on this group.

Group 5, shown on Table 32, represents the core of the Midwest. Here the tendency is for moderate positive scores on Socioeconomic, Racial-Educational, and Agriculture, and, most characteristically, high negative scores on Educational Expenses vs. Manufacturing. As has been mentioned, these states represent the manufacturing end of this latter bipolar factor.

The other three groups are presented on Table 33. All of these groups are small, the loadings low, and a pattern is hard to distinguish. Group 8 brings together three states which rationally seem to belong together, but which have surprisingly different profiles--Maryland, Delaware, and

TABLE 32
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation
Group 5

State	Factor Loading	R FACTORS Socio-Economic	Racial-Education	Educ. Expense vs. Mfg.	Frontier	Agric.
Indiana	.88	0	+	--	0	+
Ohio	.60	+	+	--	-	+
Michigan	.54	+	0	-	-	0
Wisconsin	.41	+	+	--	0	+

FACTOR SCORE LEVELS

+ = + .26 to + .99 — - = - .26 to - .99

0 = - .25 to + .25 — -- = -1.00 to -1.99

TABLE 33
Varimax Factor Loadings and R Factor Levels
Q Analysis Based on Correlation

State	Factor Loading	R FACTORS Socio-Economic	Racial-Educ.	Educ. Expense Vs. Mfg.	Frontier	Agric.
		Group 6				
Oklahoma	.70	-	0	+	-	0
Texas	.41	+	-	0	-	+
Hawaii	-.41	++	-	0	++	+
Alaska	-.48	0	-	-	++++++	-
		Group 7				
Florida	.75	+	--	0	0	0
Nevada	.66	+	0	+	++	-
		Group 8				
Maryland	.54	+	-	-	0	-
Delaware	.48	++	-	-	+	+
Virginia	.44	-	--	-	+	-

TABLE 33
Continued

FACTOR SCORE LEVELS	
++++++ = +5.00 to +5.99	0 = - .25 to + .25
++ = +1.00 to +1.99	- = - .26 to - .99
+ = + .26 to + .99	-- = -1.00 to -1.99

Virginia. What ties them together is not their contiguity, of course, but the fact that they share enough of a profile to group together, even though with only moderate factor loadings.

For these data it is clear that the groupings resulting from Q analysis based on d presented earlier are superior to those based on r. Not only are the groupings clearer, with fewer small, unclassifiable groups, but conceptionally D analysis is closer to the common research questions of the individuals grouping political units. One is most often interested in differentiating groups of nations, states, counties, or cities which have similar percentage or per capita scores, i.e., similar as measured by the DSI, rather than in similarity of profile shape, as measured by the intercorrelation matrix. Q based on r has been used many times with relative success (e.g., Russett, 1967, pp. 23-35), but if the same data were used in a D analysis, even clearer groupings might have emerged. In the study just cited, Russett did make use of distance measures in a comparative analysis (pp. 49-56). There were, however, two important differences between his use of distance and the use made in the present study. First, he used hierarchical clustering

rather than factor analysis. His objection to hierarchical clustering and the differences that resulted from the use of this method need not be cited here. Second, he took the conventional approach of using factor scores rather than raw data. This, as has been shown, can obscure many differences which are present in the data.

This is not to say that Q based on correlation does not have a contribution to make in the grouping of political units. Its contribution comes not from simple Q analysis, however, but from the approach of profile analysis, the subject of the next section.

Profile Analysis

Profile analysis takes a Q analysis based on correlation and refines it. It seeks out a group of cases with profiles of similar shape and, through an analysis of D3I, differentiates between clusters of cases (Guertin and Bailey, 1970, pp. 259-92). The question to be answered here is, how useful is this approach for distinguishing levels within shapes (patterns) in the study of political units, specifically, for these data in the study of the 50 states?

To answer this question the profile analysis program was run using the raw scores of the states on the 53 characteristics. The program distinguished six shapes, broken down into 11 patterns. These are summarized on Table 34. The labels given these patterns, as was true of the group labels, are type names rather than absolute designations.

TABLE 34
Shape Families and Patterns
Profile Analysis

Shape Family	Pattern
1. The South	1. Deep South 2. Border States
2. The Plains	1. Heartland 2. Peripheral Plains 3. Dakotas
3. Midwest-Eastern Seaboard	1. Midwest 2. Urban: Small Area, Big Population 3. Urban: Commercial Centers
4. Upper New England	
5. The West	
6. The Southwest	

As will be seen, there is some overlap of states from pattern to pattern, and for some it was difficult to find a label which would adequately express the tie between the states.

The output of the profile analysis program is lengthy; only a portion of its potential usefulness will be illustrated here. The program can establish types of cases, but beyond that it does what none of the other factor analysis programs already discussed are capable of doing--it gets back to the actual variables which determine the types. In discussing D analysis based on raw scores it was necessary to refer repeatedly to factor scores of representative states, even though it was raw variable scores which were used as data rather than factor scores. The reason is obvious: the case

of referring to five factors in contrast to 53 variables. This left open the question of how a grouping of states was determined by the actual variable scores. Just as the practice of using factor scores as data results in a loss of information, so a discussion of groups when based on factor scores rather than variable scores ignores much information. There can be, and often are, many subtle differences between groups with apparently similar factor scores which are obtainable only by examining the raw scores.

A discussion of actual patterns should make this clearer. Of the many matrices and interpretative tables which might be prepared from output from the profile analysis package, two have been chosen. One of these is a list of the states with the lowest d's when compared with each pattern. An example may be found in Table 35, the Deep South. On most of the similar tables which follow, those states which are marked are the pivot profiles, i.e., the states which were used to define the cluster (determine the pattern). Also included in these tables are other states which have small d's in comparison with the pattern. As mentioned many times, two cases may be compared either with respect to profile shape (correlation) or absolute size of scores (distance). Those states which determine a given pattern are those whose profiles are similar in shape and at the same level. In most cases, there are also other states with rather different shaped profiles but with overall similarity of scores which by some standards could be included in the group. In the

TABLE 35
Shape Family 1, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
Alabama	.05	Tennessee	.15
Georgia	.09	Arkansas	.15
North Carolina	.10	Virginia	.24
South Carolina	.10	Louisiana	.26
		Mississippi	.29

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
% population Negro	1.85
% registrants failed mental test of Selective Service	1.79
% population 21 and over voting in the presidential election	-1.75
% families with annual income less than $3000	1.66
Per capita murder	1.55
% population 25 years or older with less than 5 years of schooling	1.55
Per capita value retail sales	-1.55
% population 25 years or older who completed high school or more	-1.45
Per pupil expenditure	-1.42
Per capita personal income	-1.41
% employed persons in white-collar jobs	-1.40
% population White	-1.30
Personal income per pupil	-1.28
High school graduates in 1959-60 as % of 8th grade enrollment in 1955-56	-1.21
Pupils per classroom teacher in elementary and secondary	1.18
% 5 and 6 year olds in school	-1.15
Median school years nonwhite as % of median school years White	-1.15
% population 14-17 enrolled in school	-1.15
% of children in elementary school enrolled in private school	-1.13
Average teacher salary	-1.13

case of Table 35, six states with small d's were used to help determine the pattern. Notice that in these d scores, unlike the DSI, the smaller the value, the more similar the state is to the pattern, with 0.0 indicating perfect congruence.

The second type of information from the profile analysis output which will be utilized for each pattern--and the most enlightening--is a ranking of variables whose scores were the most important in discriminating the states of the pattern from those of all other patterns; in other words, the variables over which there was greatest agreement among the states of the pattern. This may be seen in the second half of Table 35, where is presented a listing of the variables with the highest weighted means for the Deep South. These means are in standard score form, with ±1.0 indicating one standard deviation above or below the mean. These values are the mean scores of the pivot profiles on a variable weighted in relation to the size of the d factor loading of each of the pivot profiles. Thus the pivot profile with the smallest d would carry the most weight in establishing the weighted mean on a variable. Those states which were not among the pivot profiles do not contribute toward the determination of these weighted means.

From the standpoint of interpretation, Table 35 shows what it means, within the limits of these data, to be the South. The variables with the highest weighted means are racial, educational, and economic. Present are such variables as income levels and educational achievement which would

[illegible] relatively undeveloped region. (For a factor analytic study of the South as a undeveloped region, see Hartshorn, 1971.) Note that this area is characterized by having a relatively small percentage of children attending private school (weighted mean, -1.13). This is one variable whose mean in the face of the racial crisis has no doubt shifted in the years since the 1960 census. The great number of variables with large weighted means gives some indication of the South's atypicalness in comparison with the country as a whole. This region stands as the most distinct region, the most atypical region, in the United States.

Compare the variables which determine this pattern with those which determine the related pattern designated Border States, which is shown on Table 36. While several Deep South states have small d's when compared with this pattern, only the Border States served as pivot profiles. Absent from the high ranking variables is the racial element as a vital determiner (weighted mean of percentage Negro, .76). Economically and educationally the region appears even more depressed than the Deep South. In other words, one cannot point to the "Blackness" of the South as a necessary component of the economic and educational level, but instead to other factors even more prevalent in the Border States than in the Deep South.

Shape Family 2 consists of the Plains states, and includes within it three distinct patterns. These are shown on Tables 37, 38, and 39. The patterns of all three groups of states--

TABLE 36
Shape Family 1, Pattern 2

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Arkansas	.03	*West Virginia	.35
*Kentucky	.12	South Carolina	.35
Tennessee	.20	Georgia	.35
North Carolina	.26		

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
% families with annual income less than $3000	2.10
% population 25 years and older who complete high school or more	-1.82
% population 25 years and older who complete college	-1.75
Average teacher salary	-1.72
% registrants failed mental test of Selective Service	1.64
Per pupil expenditure	-1.61
Per capita personal income	-1.58
Per capita expenditures of state and local governments for education	-1.43
Per capita value retail sales	-1.41
% employed persons in white-collar jobs	-1.40
% families with annual income $10,000 or more	-1.34
% change in public school enrollment 1950-51 to 1960-61	-1.31
Personal income per pupil	-1.30
% population 25 years or older with less than 5 years of schooling	1.28
% population urban	-1.22

TABLE 37
Shape Family Z, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d
Kansas	.03
Nebraska	.08
Iowa	.09

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Per capita value farm products	1.89
% of land area represented by all land in farms	1.85
% voted Democratic	-1.65
Pupils per classroom teacher in elementary and secondary	-1.41
% of state and local government employees who were employed as teachers	1.27
% population 65 years of age and over	1.18
Per capita motor vehicle registrations	1.13
% 5 and 6 year olds enrolled	1.12
% employed persons in agriculture	1.02
% population rural farm	1.01

TABLE 38
Family 2, Pattern 2

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Montana	.03	Oregon	.22
*Kansas	.11	Oklahoma	.26
*Idaho	.12	Minnesota	.27
*Wyoming	.12	South Dakota	.27

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Per Capita motor vehicle registrations	1.69
% of state and local government employees who were employed as teachers	1.52
Per capita value farm products	1.48
Pupils per classroom teacher in elementary and secondary	-1.13
% of civilian labor force male	1.10
% registrants failed mental test of Selective Service	-1.03
% population 14-17 enrolled in school	.97
% population 25 years and older who completed high school or more	.96
% population 25 years and older with less than 5 years schooling	-.92
% employed persons in agriculture	.91

TABLE 39
[illegible] Family 2, Pattern 3

States with Lowest d's on This Pattern
When Compared with This Pattern

States	d
North Dakota	.01
South Dakota	.04

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
% employed persons in agriculture	3.21
% population rural farm	2.74
Per capita value farm products	2.37
Pupils per classroom teacher in elementary and secondary	-2.32
% employed persons in manufacturing	-1.77
% civilian labor force male	1.74
Farm land as % of all land area	1.71
Expenditures for public elementary and secondary education as % of personal income	1.70
% population urban	-1.68
Per capita value added by manufacturing	-1.55
% employed persons in educational services	1.34

..., ...ly, states in the three patterns overlap--
...s agriculture and a high involvement in education, yet the levels are different. Compare Patterns 1 and 3. Their greatest difference is in level; they illustrate well how profile analysis can take a group of cases of essentially the same shape profile and cluster them into levels. Pattern 3 has a much deeper emphasis upon agriculture than do Patterns 1 and 2, as shown by the size of the weighted means of the first three variables in Table 39. For the Dakotas also, being less urban (-1.68) than the other regions, education is more expensive (expenditures as a percentage of personal income, 1.70), with even lower pupil-per-teacher ratios (weighted mean, -2.32) than the other patterns (-1.41 and -1.13, respectively). Pattern 2 shows less involvement in agriculture than the other two, though still sizeable, and a higher educational level. The states in Pattern 1 are strongly Republican (percentage voted Democratic, -1.68).

Shape Family 3 consists of the urban Midwest and Northeast. All three patterns have some profile points that are opposite in direction from those of Shape Family 2. In all three patterns there is an emphasis upon manufacturing rather than agriculture, relatively low educational expenses, high population density, and high percentage urban. A comparison of the three patterns on variables which measure these socioeconomic areas can be seen in Tables 40, 41, and 42.

The combination of educationally related variables high on Patterns 1 and 3, Tables 40 and 42, show some of

TABLE 40
Shape Family 3, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Pennsylvania	.03	Michigan	.18
*Ohio	.05	Indiana	.29
*Illinois	.16	*Missouri	.36
		*Wisconsin	.39

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Total public school enrollment	1.56
% of states & local government employees who were employed as teachers	-1.48
% of children in elementary school enrolled in private schools	1.38
Per capita value added by manufacturing	1.37
% employed persons in manufacturing	1.27
% total population enrolled in public school	-1.27
% employed persons in educational services	-1.20
Expenditures for public elementary & secondary education as % of personal income	-1.13
High school graduates in 1959-60 as % of 8th grade enrollment in 1955-56	1.02

TABLE 41
Shape Family 3, Pattern 2

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Massachusetts	.03	Connecticut	.19
*Rhode Island	.14	*New Jersey	.21

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Population density	3.09
% voted Democratic	2.29
% total state & local government expenditures for education	-2.20
% population foreign born	2.00
% total population enrolled in public school	-1.83
Expenditures for public elementary and secondary education as % of personal income	-1.63
% civilian labor force male	-1.60
% population urban	1.53
% of children in elementary school enrolled in private school	1.51
Personal income per pupil	1.46
% employed persons in manufacturing	1.36
Farm land as % of all land area	-1.33
% population rural nonfarm	-1.26
% of revenue for public elementary and secondary education from local sources	1.25
% of state and local government employees who were employed as teachers	-1.25
% population rural farm	-1.19
% employed persons in agriculture	-1.17
Per capita expenditures of state and local government for education	-1.17
Per capita value added by manufacturing	1.13

TABLE 42
Shape Family 3, Pattern 3

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Illinois	.09	Pennsylvania	.25
*New York	.11	Michigan	.31
Ohio	.23	*New Jersey	.38

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Per capita value selected services	3.46
Total public school enrollment	2.25
Personal income per pupil	1.86
% population foreign born	1.76
Per pupil expenditure	1.66
% of children in elementary school enrolled in private school	1.54
% total population enrolled in public school	-1.45
Average teacher salary	1.45
Per capita personal income	1.44
% population urban	1.42
% population rural nonfarm	-1.40
% 5 and 6 year olds in school	1.33
% families with annual income $10,000 or more	1.28
Per capita motor vehicle registrations	-1.25
Per capita value added by manufacturing	1.23
% employed persons in educational services	-1.23

the reason for [illegible] educational expenses. For one thing, the population size indicator, total public school enrollment, loads on these patterns. Relatively high population density has already been mentioned. This is in contrast to the Plains states, where the population density was low and school expenses high. The second main ingredient is that these areas are the ones with a higher percentage of children attending private schools, on one hand, and a smaller percentage of children enrolled in public school--an indicator of percentage of the population of school age--on the other. Hence in these Midwestern and Eastern Seaboard states there are higher concentrations of population, lower percentages of school age children, and lower percentages of these school age children in public school. Consequently, education quite naturally occupies a smaller share of the economy.

Shape Families 4, 5, and 6 have only one pattern each. Shape Family 4 is shown on Table 43. This shape includes Maine, New Hampshire, and Vermont, a grouping seen in the other analyses. Because of the R factor scores of these three states, they have been classified as both low socioeconomically and nonagricultural. In addition, they represent the "White" pole of the Racial-Educational factor. The list of variables found on Table 43 makes it possible for the first time to determine just what variables have grouped these states together. The generalization about their being both low socioeconomically and nonagricultural seems to hold, as indicated by the weighted means of the variables

TABLE 43
Shape Family 4, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d
Maine	.05
New Hampshire	.07
Vermont	.11

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
Median school years nonWhite as % of median school years White	1.75
% of revenue for public elementary and secondary education from local sources	1.41
% population rural nonfarm	1.39
Farm land as % of all land area	-1.16
Per Capita burglary	-1.12
Per capita murder	-1.10
% population 65 years of age and over	1.10
% employed persons in manufacturing	1.06

percentage rural nonfarm (1.39) and farm land as a percentage of all land area (-1.16). Percentage of the population urban, not in the table, has a weighted mean of -0.74. Note that one area which makes these states both nonurban and non-South, so to speak, is their small percentage of crime, with per capita burglary (high on Socioeconomic because of the close relationship to urban), -1.12, and per capita murder (high in the South), -1.10.

As for Factor 2, Racial-Educational, Upper New England does represent the White pole of this bipolar factor, but not so much in percentage of White (weighted mean, 0.78) and Black (-0.83) as in other related variables. The variable with the highest mean (1.75) is median school years nonwhite as a percentage of median school years White. In other words, in comparison with the other states and with Whites within these three states, the nonwhites tend to have been in school longer. New Hampshire, for instance, was the only one of the 50 states which had a positive score on this variable, indicating that in this state the nonwhites (includes "other races," i.e., Orientals, etc., as well as Negroes) are better educated on the whole than the Whites. The low per capita murder rate, high in the South, has already been mentioned, and is a second variable which associates Upper New England with the White pole. A third variable which puts these states in juxtaposition to the South is the percentage of revenue for public elementary and secondary education from local sources. For Upper New England the

weighted mean for this variable is 1.41; for the South, Shape 1, Pattern 1, -1.04. Other patterns have positive means on this variable, particularly Shape 3, Pattern 1, where the rest of New England loads, but none so high as that of Upper New England. No pattern other than the South has a sizeable negative mean. These means show that in the South more of the expenditures for education tend to come from the state government whereas in states like Maine, New Hampshire, and Vermont more comes from local sources. This characteristic can be viewed as a concomitant variable indicating a different life style rather than being necessarily tied to either the racial balance or the quality of education in each region.

Shape families 5 and 6, Tables 44 and 45, may be discussed together. Both shapes show that for these data there are few socioeconomic factors which separate the regions of the West and the Southwest from the rest of the country. Most of the variables listed on Tables 44 and 45 are educationally related variables. Both Shape Families 5 and 6 have a higher level of education than the other shape families and both spend more for education. Shape Family 6 differs from all the other patterns in having a young population (percentage of the population 65 years of age and over, -1.15), a high birth rate (1.83), and thus a larger percentage of children to educate (percentage of the total population enrolled in public school, 1.77). As a result,

TABLE 44
Shape Family 5, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Washington	.03	*Minnesota	.17
*Oregon	.04	Oklahoma	.28
*Colorado	.12	Wyoming	.29

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
% population 14-17 enrolled in school	1.16
Per capita expenditures of state and local government for education	1.12
% population 25 years or older who completed high school or more	1.09
% population 25 years or older who completed college	1.07
% registrants who failed mental test of Selected Service	-.98
% employed persons in white-collar jobs	.92
% population 25 years or older with less than 5 years of schooling	-.90
Per capita motor vehicle registrations	.88
% employed persons in educational services	.87
% of state and local government employees who were employed as teachers	.83

TABLE 45
Shape Family 6, Pattern 1

States with Lowest d's on This Pattern
When Compared with This Pattern

State	d	State	d
*Utah	.02	Washington	.36
Colorado	.27	Oregon	.37
*New Mexico	.31	*Arizona	.38

*Pivot Profiles

Variables with Highest Means
Weighted According to Discriminating Ability in Pattern

Variable	Weighted Mean
% total state and local expenditures for education	2.42
% employed persons in educational services	2.39
Expenditures for public elementary and secondary education from local sources	1.89
Birth rate	1.83
% total population enrolled in public school	1.77
% population 5 years and older migrant from a different county	1.66
% population 25 years and older who completed college	1.64
Per capita expenditures of state and local government for education	1.55
% population 65 years of age and over	-1.15
% of children in elementary school enrolled in private school	-1.12

a larger part of the total state and local budgets goes toward education (2.42).

In discussing these patterns it may be noted that some states have not been grouped on any of the patterns. How should these states be grouped? Among the voluminous output of the profile analysis package is a matrix of d scores for each case on each pattern, using variable weightings and standard scores. From this matrix one can determine which pattern is closest to each case's scores in terms of distance. Table 46 gives the eight states which heretofore have not been placed in any of the patterns, together

TABLE 46
Unclassified States and Their Closest Patterns

State	Closest Pattern	Distance
Texas	3-1 Midwest	0.63
California	3-3 Urban: Commercial Centers	0.69
Delaware	5-1 West	0.41
Maryland	5-1 West	0.42
Florida	5-1 West	0.89
Nevada	6-1 Southwest	1.77
Hawaii	6-1 Southwest	2.56
Alaska	6-1 Southwest	15.97

with the patterns in which they fall. There are two ways to treat these states. One would be to add them to the pattern groups to which they are closest. If this were done, some of the pattern labels would have to be changed. For instance, Shape Family 5, Pattern 1, the West, becomes something else when Delaware, Maryland, and Florida are

added, and one would need to go to the variables to determine a label. Another way to treat these states, especially those whose smallest d is much larger than those of others within a given pattern, would be to admit simply that these states are unclassifiable. The profile analysis program presents patterns only when two cases are loaded heavily, or three moderately, on a pattern. This means that when cases have profile shapes different from those of all other cases, they will fit into no shape family. If one is dealing with shape, it is probably better to state, "Alaska has a shape unlike that of any other group of states, with standardized d's ranging from 15.97 (Southwest) to 20.92 (Urban: Commercial Centers)," than to place Alaska with the Southwest.

One more use of the output provided by this program may be cited. Through manual graphing one can compare individual cases with a given pattern. An example of this is a comparison of Mississippi with the pattern, the South. The question might be asked why Mississippi has a d as large as 0.29 on this pattern. To answer this, first a graph was made of the pivot profiles with the lowest d's on this pattern, using as data points their standard scores on the variables with the highest weighted means, or, in other words, the variables which were most responsible for distinguishing this pattern from all other patterns. The profiles formed, as expected, a tight cluster. When a graph of Mississippi's scores on the same variables was superimposed upon this cluster, it could be seen that on the variables

which discriminate the South from the other patterns. Mississippi had scores which were more extreme than those of the cluster. Mississippi's profile stands not between the South and the mean of the states, but beyond the South, with a steeper profile on these variables. Mississippi is, so to speak, extreme South--deeper than the Deep South.

Profile analysis has great potential usefulness which needs to be explored in many more studies. It is best used, not as a quick, overall grouping procedure, but as a procedure which enables one to make an in-depth investigation of the meaning of types. Specifically, profile analysis enables one to describe such groupings as the South, Upper New England, the Plains, and do so in terms of the actual variables which explain the differences; moreover, profile analysis helps the researcher to go beyond the process of grouping to explain why cases group together. The strength of this approach lies in heightened understanding beyond that possible in most grouping procedures.

CHAPTER 6

CONCLUSIONS

The purpose of this study was to present a sample of factor analytic comparisons of methods of grouping political units on characteristics. The data selected were 53 socio-economic, educational, and political characteristics of the 50 states for the year 1960.

Factor analysis has been used most frequently in determining dimensions or factors within the intercorrelations of characteristics. Much less emphasis has been given to the grouping of cases. Where cases have been grouped factor analytically, there has been no general agreement on the method to be used. Indeed, many researchers seem to be unaware that there are several alternatives in applying factor analysis to grouping but have tended to follow the method most often used in their respective disciplines. Little work has been done in cross-method comparisons.

Before transposing the data an R-type analysis was made to determine the dimensions within the characteristics. The five interpretable factors were labeled Socioeconomic, Racial-Educational, Educational Expenses vs. Manufacturing, Frontier, and Agriculture. The first factor corresponds to the largest factor found in most studies of political units. Factor 2 is related to racial or educational factors

found in the two studies which have been made in the area of educational dimensions of the United States, but neither includes as wide a range of variables. The combinations of variables which load on Factors 3 and 4 seem to be unique to this study because some of the variables themselves have not been widely used.

In factor analyzing cases two general type of matrices may be used, the intercorrelation matrix and a matrix based on distance measures. The first focuses on profile shape and the second on distances between scores. This study was intent upon both exploring each of these two general approaches in depth and of comparing these two methods with each other. In focusing on distance measures, an attempt was made to answer the question of what differences in grouping results from using as data the standardized raw scores or factor scores. Factor scores weighted according to three separate criteria were used. Factor scores were given, successively, unit weights, weights based on the percentage of variance accounted for by each factor, and weights based on the percentage of educational variables loading on each factor.

The purpose of focusing on the grouping of cases on the basis of the intercorrelation matrix was to compare the groupings obtained by this method with those based on distance, and to illustrate how profile analysis, a refinement of the correlation approach, could be used to relate the variables to types of cases.

A comparison of the groupings formed by the use of distance measures for the four types of data employed showed that for these data no great differences between the groupings could be discerned. However, a comparison of the use of different data in the grouping of Florida counties showed quite different results. In the Florida study there were great differences in groupings when using standardized raw scores and using unit weighted factor scores. The factor structure of the 50 states seemed to be sufficiently stable to result in similar groups for the four distance analyses. In the analyses of the Florida data, however, the R factors were so greatly different in percentage of variance each explained that equating them in the use of factor scores brought very different results.

Some of the groupings which resulted from factor analyzing the intercorrelation matrix were in general similar to those based on the DSI (distance similarity index), but there were some states which were unclassifiable. Several groups had to be split into two, the group of states loading positively and those loading negatively. This further complicated the interpretation. Empirically, grouping cases on correlation resulted in less satisfactory results than grouping cases on distance, and, conceptually, focusing on profile shape is less desirable in grouping political units than focusing on score magnitude.

Profile analysis, on the other hand, yielded information not available from any of the other analyses. From this

type of analysis were determined which variables explained various groups of states. Six shape family types were included: the South, the Plains, Midwest-Eastern Seaboard, Upper New England, the West, and the Southwest. The first three shape families were further broken down into several patterns. Profile analysis is valuable to use when one is interested in doing more than establishing groups, for it goes beyond the other grouping procedures to help to explain in variable terms why cases group as they do.

Several conclusions may be reached as a result of this study. First, in factor analytic grouping of cases much more thought needs to be given to determining the general approach, correlation or distance. In the grouping of political units, magnitude, i.e., distance, is usually more important than profile shape. Second, the indiscriminate use of factor scores as data in distance analysis or in any other type of analysis, should be avoided. Consideration should be given to the possible effects of different weightings given factor scores, especially when the percentages of variance explained by each factor vary widely. The sharp contrast between the analyses of the United States and Florida data underscores the importance of this and suggests that the point deserves further testing. Sometimes the research question will suggest a weighting criterion that can be applied to the factor scores. If, however, the research question is, what groupings emerge from a particular set of variables, the best practice would be to use standardized

raw scores. Even when factor scores are used for some purpose, it would be advisable to compare the results with those obtained from the use of raw scores. In any event, factor scores should be employed deliberately, with cognizance of the relationship of their use to the research question, rather than used as a shortcut to grouping.

Third, rather than taking a one-shot approach to grouping, with one method used to the exclusion of others, much more can be learned from a comparative approach. This has already been suggested in referring to the use of raw scores versus factor scores. Of particular worth for this purpose is the approach of profile analysis. Combined with a distance analysis, profile analysis shows much promise for in-depth understanding of groups of cases.

APPENDIX

Variables Used in the 50 States Study*

Economic Characteristics

% families with annual income less than $3000 (2:1)
% families with annual income $10,000 or more (2:1)
Per capita personal income (5:431)
Per capita value farm products (Computed from 5:7 and 5:850)
Per capita motor vehicle registrations (Computed from 5:7 and 5:766)
Per capita value added by manufacturing (Computed from 5:7 and 5:1094)
Per capita value retail sales (Computed from 5:7 and 5:1167)
Per capita value selected services (Computed from 5:7 and 5:1181)
Personal income per pupil (1:40)
% employed persons in agriculture (Computed from 2:1)
% employed persons in manufacturing (Computed from 2:1)
**% employed persons in educational services (Computed from 2:1)
% employed persons in white-collar jobs (Computed from 2:1)
% civilian labor force male (2:1)
**% of state and local government employees employed as teachers (5:567)
Farm land as % of all land area (Computed from 5:840 and 5:222)

Population Characteristics

Population density (2:1)
% population 65 years of age and over (2:1)
% population foreign born (2:1)
% population White (4:56)
% population Negro (4:56)
% population other races (4:56)
% population urban (4:107)
% population rural nonfarm (4:107)
% population rural farm (4:107)
Males per 100 females (5:17)

Political Characteristics

% population 21 and over voting in the 1960 presidential election (5:535)
% presidential vote Democratic (5:484)

*Two numbers are given after each variable. The first identifies the source (see list of sources at the end), the second the table within that source from which the data series was taken.
**Educational Variables.

Variables Used in the 50 States Study
Continued

Crime Indices

Per capita murder (5:190)
Per capita burglary (5:190)

Population Growth and Migration

% population increase, 1950-60 (2:1)
Birth rate (5:53)
**% change in public school enrollment, 1950-51 to 1960-61 (1:11)
% population 5 years and older migrant from a different county (2:1)

Educational Characteristics

**Median school years nonwhite as a % of median school years White (Computed from 3:6)
**% registrants failed mental test of the Selective Service (5:335)
**% population 25 years and older who completed college (Computed from 5:149)
**High school graduates in 1950-60 as % of 8th grade enrollment in 1955-56 (1:35)
**% 5 and 6 year olds in school (4:114)
**Pupils per classroom teacher in elementary and secondary (1:28)
**% population 25 years and older with less than 5 years of schooling (2:1)
**% population 25 years and older who completed high school or more (2:1)
**% of children in elementary school enrolled in private school (4:105)
**% population 14-17 enrolled in school (4:105)
**% total population enrolled in public school (1:15)
**Average size school district (Computed from 3:1)
**Total public school enrollment (3:20)

Educational Support

**Per pupil expenditure (5:53)
**Average teacher salary (5:162)
**Per capita expenditures of state and local government for education (5:423)
**Expenditures for public elementary and secondary education as % of personal income (3:37)
**% of revenue for public elementary and secondary education from local sources (3:32)
**% total state and local expenditures for education (5:553)

**Educational Variables

Variables Used in the 50 States Study
Continued

Sources

1. Research Division, National Educational Association. Rankings of the states, 1962. Research Report 1962-R1. Washington, D. C.: 1962.

2. U. S. Department of Commerce, Bureau of the Census. County and city data book. Washington, D. C.: Government Printing Office, 1962.

3. U. S. Department of Health, Education, and Welfare, Office of Education. Digest of educational statistics. Washington, D. C.: Government Printing Office, 1962.

4. U. S. Department of Commerce, Bureau of the Census. General population characteristics, vol. 1, U. S. summary. Washington, D. C.: Government Printing Office, 1962.

5. U. S. Department of Commerce, Bureau of the Census. Statistical abstract of the United States, 1962. Washington, D. C.: Government Printing Office, 1962.

Variables Used in the Florida Study

Socioeconomic Characteristics

Per capita personal income
% families with annual income $10,000 or more
% families with annual income less than $3000
Personal income per pupil
Welfare recipients per 1000 population
Per capita hotel and motel units
Per capita mobile homes
Per capita value farm products
Per capita value retail trde
Per capita value selected services
Per capita tourist trade, as measured by county destination
    of incoming automobile tourists
Non-exempt assessed valuation per pupil
% of personal income received in agriculture
% of personal income received in manufacturing

Population Characteristics

% population 65 years and older
% population under 18
% population urban
% population rural nonfarm
% population rural farm
% population foreign born
% population nonwhite
Population density

Employment Characteristics

% employed persons in white-collar jobs
% employed persons in educational services
% employed persons in agriculture
% employed persons in manufacturing
% employed persons in retail sales
% males in labor force

Political Characteristics

% population voting in 1968 presidential election
% registered voters Negro
% voting for Richard Nixon, 1968
% voting for Hubert Humphrey, 1968
% voting for George Wallace, 1968

Variables Used in the Florida Study
Continued

Population Growth and Migration

Birth rate
% native population born in Florida
% population moved into county between 1955 and 1960
% school enrollment increase, 1957-58 to 1967-68
% enrollment transfers from out-of-state
% change in school membership first month 1968-69 over first month 1967-68
School withdrawals as % of enrollment
Ratio of number housing units built from 1950 to 1960 to number built before 1950
% population increase 1950-60
% increase of population 65 and over 1950-60

Economic Growth

% change in per capita income from 1960 to 1967
Change in Index of Taxpaying Ability from 1958-59 to 1968-69
% increase in current expenditures per pupil
% increase in local school revenue from 1957-58 to 1967-68
% increase in annual salary paid instructional staff 1957-58 to 1967-68

Educational Characteristics

% population 25 years and older with less than 5 years of schooling
% population 25 years and older who completed high school or more
Median school years completed by population 25 years and older
% population 14-17 enrolled in school
% population 20-24 enrolled in school
% 1968 high school graduates who entered college in 1968-69

School Characteristics

Pupils per teacher, elementary
Pupils per teacher, 7-12
% instructional personnel below Rank III
% instructional personnel Rank II or above
% instructional personnel on Continuing Contract
% elementary teachers taught in field
% major high school classes taught by teachers in field
% teachers men
High school enrollment as % of elementary
Teaching positions as % of academic units earned.

Variables Used in the Florida Study
Continued

Sources

Bureau of Economic and Business Research, College of Business Administration, University of Florida. Florida statistical abstract, 1969. Gainesville, 1969.

State of Florida, Department of Education. Research report no. 73, ranking of the counties . . . 1968. Tallahassee, 1969.

State of Florida, Secretary of State. Tabulation of official votes cast in the general election, November 5, 1968. Tallahassee, 1968.

U. S. Bureau of the Census. County and city data book. Washington, D. C.: U. S. Government Printing Office, 1962.

TABLE A.1
Varimax Rotated Factor Loadings
R Analysis

Characteristic	Socio-Economic 1	Racial Educ. 2	Educ. Expense Vs. Mfg. 3	Frontier 4	Agric. 5
Average teacher salary	.85	.14	-.07	.28	-.31*
% population urban	.85	.07	-.12	-.26	-.26
Per capita personal income	.79	.38*	-.16	.28	-.14
% families with annual income $10.000 or more	.78	.14	-.11	.50	-.20
% employed persons in white-collar jobs	.77	.25	.13	.24	-.25
Per pupil expenditure	.77	.37*	.03	.30*	-.07
% population rural nonfarm	-.77	-.10	-.00	.46	-.14
Per capita value selected services	.75	.15	-.15	-.12	-.03
Personal income per pupil	.71	.39*	-.40	.15	-.16
% families with annual income less than $3000	-.69	-.58	.03	-.21	.28
% population 25 years and over who completed college	.69	.10	.34*	.24	-.32*
% 5 and 6 year olds enrolled in school	.66	.41	-.18	.02	.10
% population foreign born	.63	.36*	-.17	-.02	-.29
Per capita burglary	.53	-.21	.14	.03	-.14
Total public school enrollment	.48	-.21	-.43	-.32*	.14
% of children in elementary school enrolled in private school	.48	.43	-.46	-.14	.05

Note: On this and all subsequent factor tables the factors are listed in the order of percentage of variance accounted for by each. Loadings .40 and above are underlined; those from .30 to .40 are marked with an asterisk. All tables are based on the U. S. data unless marked otherwise.

TABLE A.1
Continued

Characteristic	Socio-Economic 1	Racial Educ. 2	Educ. Expense Vs. Mfg. 3	Frontier 4	Agric. 5
% population 21 and over voting in 1960 presidential election	.23	.91	.02	-.03	-.01
% population Negro	-.20	-.89	-.19	-.16	.02
% population 25 years and over with less than 5 years of schooling	-.29	-.89	-.11	-.06	-.08
Per capita murder	-.21	-.87	-.02	.19	-.05
% registrants failed mental test of Selective Service	-.32*	-.85	-.23	-.12	-.11
Median school years nonwhite as % of median school years white	-.04	.82	-.25	-.33*	-.06
% population white	.05	.76	.08	-.22	-.03
% of revenue for public elementary and secondary schools from local sources	.12	.64	-.33*	-.25	.14
Per capita value retail sales	.55	.62	.14	-.05	.13
Pupils per classroom teacher	-.14	-.38	-.31*	-.17	-.26
High school graduates in 1959-60 as % of 8th grade enrollment in 1955-56	.54	.57	-.02	-.24	.21
% population 25 years and older who completed high school or more	.47	.55	.44	.36*	-.05
% population 14-17 enrolled in school	.38*	.54	.44	.09	.32*
% of state and local government employees who are teachers	-.08	.22	.81	.21	.18
Expenditures for public elementary and secondary schools as % of personal income	.02	-.14	.81	.01	.24
% employed persons in manufacturing	.13	.07	-.74	-.36*	-.36*
% employed persons in educational services	-.17	.31*	.71	.18	-.01

TABLE A.1
Continued

Characteristic	Socio Economic 1	Racial Educ. 2	Educ. Expense Vs. Mfg. 3	Frontier 4	Agric. 5
Per capita value added by manufacturing	.37*	.21	-.70	-.31*	-.23
Per capita motor vehicle registrations	.09	.39*	.67	.02	.27
% total population enrolled in public school	-.45	-.38*	.60	-.03	.09
Per capita state and local expenditures for education	.49	.29	.58	.31*	.05
% civilian labor force male	-.18	.32*	.42	-.15	.40
% total state and local expenditures for education	.03	-.06	.27	.11	.04
Average size school district	.06	-.08	.03	.92	-.14
Males per 100 females	.04	.06	.27	.90	.02
% population 5 years and older migrant from a different county	.16	-.07	.48	.68	-.05
% change in public school enrollment, 1950-51 to 1960-61	.51	-.02	.12	.61	-.21
Birth rate	.03	-.17	.46	.39	.05
% population 65 years of age and over	-.13	.52	-.24	-.58	.26
% population increase, 1950-60	.50	-.17	.19	.55	-.19
% population other races	.13	-.07	.10	.46	.03
Farm land as % of all land area	-.07	-.00	.07	-.15	.80
Per capita value farm products	-.23	.32*	.49	-.02	.70
% employed persons in agriculture	-.47	.02	.41	-.07	.70
% population rural farm	-.36	-.00	.23	-.13	.69
% 1960 presidential vote Democratic	.24	-.27	-.24	-.08	-.54
Population density	.41	.05	-.37*	-.24	-.41

TABLE A.2
Simple Loadings Primary Factors
Q Analysis Based on Raw Score Distances

State	1	2	3	4	5	6	7
South Carolina	1.00	-.09	.03	.03	-.03	-.08	-.02
Alabama	.98	-.02	.08	-.00	-.00	-.02	-.03
Georgia	.92	-.02	.15	-.06	.11	-.02	-.07
North Carolina	.89	-.04	.02	.10	-.03	.01	.10
Mississippi	.88	.03	-.13	.03	-.14	.00	-.05
Arkansas	.81	.18	-.04	-.02	-.09	.01	.14
Tennessee	.76	.05	-.00	-.01	.05	.13	.19
Louisiana	.70	.06	.10	.04	.20	.04	-.06
Virginia	.63	.03	.05	.00	.29	.13	.10
Kentucky	.56	.17	-.05	-.09	.04	.16	.38*
Texas	.44	.16	.01	.06	.31*	.30*	-.02
South Dakota	.06	.99	-.03	-.04	-.01	-.08	-.00
North Dakota	.06	.96	-.01	-.00	-.03	-.13	-.02
Nebraska	-.06	.92	.02	-.05	-.00	.18	.02
Iowa	-.04	.84	.03	.01	-.07	.22	.10
Montana	.04	.68	.09	.23	.16	-.08	.05
Kansas	.04	.61	-.01	.27	.06	.15	.07
Idaho	.11	.59	.03	.21	.12	-.09	.17
Wyoming	.00	.51	.03	.49	.14	-.14	-.07
Minnesota	.00	.43	.21	.32*	-.02	.20	.09
Oklahoma	.30*	.33*	-.06	.22	.15	.21	.14
Massachusetts	-.00	.05	.93	.01	.00	-.07	.09
Rhode Island	.08	.01	.87	-.10	.02	-.08	.18
New Jersey	.03	.02	.82	.09	.00	.12	-.09
Connecticut	.00	-.01	.80	.21	.03	.03	.02
New York	-.02	.02	.67	.06	-.01	.25	-.14
Pennsylvania	.12	.01	.46	.07	-.03	.39*	.23
Maryland	.24	.03	.39*	.07	.32	.22	.03
Delaware	.09	.04	.34*	.25	.25	.21	-.03
Utah	.08	.05	-.05	.70	.15	.01	.08
Washington	.04	.06	.21	.70	.05	.10	.09
Oregon	.03	.18	.19	.66	.02	.02	.13
Colorado	.00	.22	.09	.54	.23	.10	.00
California	-.02	-.01	.26	.36*	.21	.36*	-.17

TABLE A.2
Continued

State	1	2	3	4	5	6	7
Alaska	-.09	-.08	-.08	.05	.65	-.05	.07
Nevada	-.04	.16	.11	.10	.62	.04	-.07
Arizona	.18	.13	-.05	.31*	.32	.11	-.08
Florida	.29	.12	.14	-.08	.31	.21	-.06
New Mexico	.27	.13	-.07	.38*	.42	-.02	-.03
Hawaii	.13	.06	.17	-.03	.42	.01	.01
Ohio	.10	.07	.30*	.13	.03	.51	.17
Illinois	.02	.12	.48	.03	.03	.49	-.03
Indiana	.15	.10	.05	.28	.03	.48	.25
Missouri	.20	.22	.19	-.05	.11	.42	.25
Michigan	.11	-.01	.30	.36	.03	.41	.05
Wisconsin	-.02	.30*	.33*	.11	-.05	.33*	.20
Maine	.11	.18	.23	.07	.10	.04	.58
New Hampshire	-.01	.10	.48	.04	.09	.04	.53
Vermont	.06	.33*	.14	.16	.08	.02	.49
West Virginia	.47	.01	-.01	.19	-.03	.04	.48

TABLE A.3
Maxplane Primary Factors
Q Analysis Based on Raw Score Distances

State	1	2	3	4	5	6	7
South Carolina	1.11	-.21	.04	.01	-.08	-.05	-.06
Alabama	1.08	-.09	.09	-.03	-.11	.02	-.03
Georgia	1.01	-.09	.28	-.10	-.18	.05	.09
Mississippi	1.00	-.01	-.17	.03	-.12	.09	-.17
North Carolina	.97	-.07	-.02	.07	.09	-.08	-.07
Arkansas	.88	-.01	-.07	-.10	.18	.11	-.11
Tennessee	.80	.16	-.09	-.12	.21	-.03	.04
Louisiana	.76	.05	.09	.03	-.18	.09	.18
Virginia	.56	.18	-.03	-.07	.06	-.02	.29
Kentucky	.36	.20	-.17	-.27	.54	-.01	.06
Illinois	-.10	.88	.40	.03	-.38*	.09	-.01
Ohio	-.01	.84	.14	.11	-.04	-.07	-.01
Indiana	.08	.78	-.18	.28	.15	-.13	-.03
Missouri	.11	.71	.05	-.18	.17	.07	.11
Michigan	.03	.67	.16	.46	-.20	-.13	-.06
California	-.08	.63	.16	.52	-.53	-.05	.13
Pennsylvania	-.00	.61	.38*	.01	.03	-.10	-.08
Wisconsin	-.13	.59	.26	.06	.12	.14	-.09
Texas	.45	.55	-.12	.04	-.21	.11	.30*
Minnesota	-.06	.41	.17	.38*	.00	.26	-.08
Oklahoma	.30*	.38*	-.18	.21	.12	.15	.12
Delaware	.04	.34*	.31*	.33*	-.23	-.00	.19
Massachusetts	-.13	-.18	1.14	-.01	.02	.10	-.05
Rhode Island	-.04	-.21	1.06	-.20	.18	.06	-.00
New Jersey	-.09	.20	.97	.14	-.37*	.10	-.07
Connecticut	-.11	.01	.94	.28	-.15	.01	-.05
New York	-.13	.45	.73	.11	-.48	.09	-.07
Maryland	.18	.34*	.37*	.06	-.14	.01	.29
Utah	.10	-.04	-.14	.92	.09	-.15	.04
Washington	.01	.12	.15	.91	.01	-.13	-.08
Oregon	.01	.00	.16	.85	.14	-.03	-.10
Colorado	-.01	.19	.03	.71	-.08	.06	.14
Wyoming	.03	-.15	.08	.65	-.06	.38*	.07
New Mexico	.33*	-.04	-.12	.49	-.06	.04	.38*

TABLE A.3
Continued

State	1	2	3	4	5	6	7
Maine	.03	-.05	.19	-.10	.89	-.07	.09
Vermont	-.01	-.03	.11	.04	.78	.06	.07
New Hampshire	-.14	-.05	.50	-.12	.75	-.08	.08
West Virginia	.47	-.08	-.11	.10	.73	-.23	-.06
South Dakota	.07	.09	.07	-.12	.08	.88	.01
North Dakota	.08	-.01	.11	-.06	.08	.85	-.03
Nebraska	-.11	.54	.03	-.14	-.01	.77	.01
Iowa	-.09	.57	.00	-.08	.08	.67	-.07
Montana	.04	-.03	.15	.25	.12	.53	.13
Wyoming	.02	.37*	-.06	.30*	.05	.42	.03
Idaho	.11	-.08	.06	.19	.33*	.41	.10
Alaska	-.08	-.14	-.12	.03	.16	-.11	.69
Nevada	-.06	.09	.11	.11	-.13	.15	.63
Florida	.28	.39*	.09	-.14	-.24	.14	.53
Arizona	.20	.21	-.14	.40	-.20	.06	.49
Hawaii	.11	.01	.19	-.07	-.02	.08	.43

TABLE A.9
Oblimin Primary Factors
Q Analysis Based on Raw Score Distances

State	1	2	3	4	5	6	7
New Jersey	.97	-.02	-.01	.06	.12	-.02	-.12
Massachusetts	.95	-.01	.02	.03	.26	-.04	.03
Connecticut	.92	-.03	.01	.30*	.36*	.03	-.01
New York	.88	-.06	-.03	-.06	-.09	-.03	-.15
Rhode Island	.87	.10	-.04	-.15	.11	-.03	.09
Illinois	.87	.03	.07	-.23	-.43	-.01	-.06
Pennsylvania	.81	.17	.02	-.05	-.22	-.08	.18
Ohio	.73	.15	.09	-.04	-.34*	-.01	.13
Michigan	.68	.12	.07	.39*	.06	.05	.07
Wisconsin	.64	.04	.32*	-.01	-.18	-.09	.15
New Hampshire	.62	.09	.15	.02	.07	.03	.41
Maryland	.62	.26	.02	-.11	-.07	.30*	-.01
California	.59	-.05	.04	.33*	.04	.25	-.13
Delaware	.57	.09	.08	.21	.07	.27	-.03
Missouri	.54	.29	.20	-.35*	-.53	.05	.17
Indiana	.48	.22	.19	.23	-.16	.02	.23
South Carolina	-.03	.99	-.08	.10	.16	-.02	-.05
Alabama	.06	.98	-.03	.00	.05	.00	-.07
Georgia	.12	.92	-.05	-.14	-.04	.11	-.11
North Carolina	.05	.90	.01	.18	.15	-.12	.06
Mississippi	-.15	.88	.10	.08	.06	-.13	-.07
Arkansas	-.03	.85	.20	-.03	-.02	-.10	.08
Tennessee	.11	.82	.07	-.10	-.16	.03	.13
Louisiana	.14	.71	.06	-.05	-.04	.21	-.09
Virginia	.18	.68	.03	-.16	-.22	.28	.05
Kentucky	.10	.67	.18	-.24	-.30*	-.01	.29
West Virginia	.11	.57	.12	.35*	.24	-.04	.42
Texas	.25	.48	.16	-.20	-.37*	.30*	-.06
South Dakota	-.06	.13	.99	-.23	-.12	-.02	-.09
North Dakota	-.07	.12	.96	-.13	-.00	-.04	-.09
Nebraska	.18	.01	.89	-.35*	-.40	-.04	-.06
Iowa	.23	.04	.84	-.24	-.33*	-.11	.01
Montana	.11	.09	.75	.23	.24	.19	-.00
Kansas	.18	.09	.69	.23	.06	.08	.03
Idaho	.06	.18	.68	.24	.24	.13	.11
Minnesota	.45	.04	.52	.36*	.17	-.00	.06
Vermont	.28	.18	.43	.20	.16	.06	.41
Oklahoma	.16	.37*	.40	.16	-.05	.16	.10

TABLE A.4
Continued

State	1	2	3	4	5	6	7
Utah	.11	.08	.28	1.11	.81	.26	.13
Washington	.44	.03	.27	1.08	.78	.14	.13
Oregon	.37*	.04	.39*	1.05	.81	.10	.16
Colorado	.29	.01	.38*	.72	.48	.30*	.03
Wyoming	.04	-.00	.66	.72	.84	.23	-.06
Florida	.31*	.33*	.07	-.44	-.48	.48	-.11
Alaska	-.06	-.04	-.05	-.10	-.10	.67	.05
Nevada	.20	-.01	.17	-.12	-.13	.84	-.09
Arizona	.11	.20	.21	.23	.06	.57	-.07
New Mexico	.00	.28	.25	.46	.33*	.49	-.02
Hawaii	.22	.15	.04	-.22	-.16	.41	-.03
Maine	.38	.24	.26	.07	.06	.05	.47

TABLE A.5
Simple Loadings Primary Factors
Q Analysis Based on Unit Weighted Factor Score Distances

State	1	2	3	4	5	6	7
South Carolina	.94	-.01	-.05	-.04	-.00	.09	.05
Alabama	.93	.03	.01	-.02	.06	.05	-.03
Georgia	.90	.05	-.01	-.03	.09	.05	.03
Mississippi	.84	-.02	.03	.18	-.01	-.15	-.05
North Carolina	.77	-.04	.15	.04	-.06	.17	.09
Arkansas	.76	-.00	.15	.18	-.13	.12	-.04
Tennessee	.69	.01	.24	.05	-.05	.20	-.01
Louisiana	.67	.33*	-.10	-.06	.22	-.12	-.01
Virginia	.65	.03	.12	.01	.09	.12	.22
Texas	.50	.25	.16	.11	.24	-.09	.01
Florida	.50	.23	.07	.04	.27	-.03	.19
Kentucky	.56	-.09	.32*	.14	-.16	.37*	.07
Utah	.02	.79	-.00	-.02	.05	.22	-.10
Wyoming	.01	.78	-.02	.22	-.04	-.04	.10
Colorado	.01	.75	.17	.10	.08	.05	.02
New Mexico	.28	.75	-.07	-.05	.04	-.03	.14
Oregon	-.03	.69	.19	.06	.05	.23	.04
Arizona	.26	.64	-.01	.03	.17	-.13	.16
Washington	-.04	.62	.22	.05	.16	.20	.04
Montana	.03	.50	.07	.44	-.08	.10	.14
Okhahoma	.30*	.39*	.24	.20	.01	.10	-.05
Indiana	.08	.09	.83	.02	.01	.07	.07
Missouri	.16	.02	.82	.08	-.03	.03	.01
Wisconsin	-.00	.09	.81	.11	.02	.04	.08
Ohio	.07	.07	.75	-.03	.18	.10	.04
Pennsylvania	.10	.03	.65	-.07	.23	.22	.02
Illinois	.06	-.10	.60	.13	.37*	-.11	.07
Minnesota	.01	.36*	.37*	.36*	.10	-.00	-.02
South Dakota	.07	.01	-.11	.96	.02	.06	.03
North Dakota	.06	.08	-.15	.94	.02	.04	.04
Nebraska	.02	-.06	.20	.84	.08	-.02	-.02
Iowa	-.00	-.03	.27	.78	.06	.02	-.06
Idaho	.03	.38*	.08	.46	-.14	.30*	.09
Wyoming	.05	.40	.25	.46	-.00	.04	-.01

TABLE A.5
Continued

State	1	2	3	4	5	6	7
New York	.08	.00	.18	.11	.75	-.01	.02
New Jersey	.09	.05	.13	.06	.73	.12	.04
California	.04	.19	.17	.20	.60	-.12	.06
Connecticut	.01	.14	.14	.01	.60	.30*	.08
Massachusetts	.02	.23	.01	-.02	.53	.51	-.08
Delaware	.12	.00	.27	.15	.47	-.07	.34*
Maryland	.28	.09	.16	.02	.46	.08	.24
Michigan	.11	.13	.42	.07	.42	.07	.05
Maine	.06	.04	.06	.12	-.02	.81	.08
New Hampshire	-.01	.03	.12	.06	.11	.78	.10
West Virginia	.26	.03	-.02	.04	-.04	.76	.05
Vermont	.03	.17	.13	.21	-.07	.63	.11
Rhode Island	.10	.13	.01	-.05	.49	.52	-.10
Alaska	-.05	-.05	-.01	-.06	-.07	.03	.66
Nevada	.06	.31*	-.02	.15	.18	.03	.59
Hawaii	.14	.11	.09	.21	.28	-.08	.52

TABLE A.6
Simple Loadings Primary Factors
Q Analysis Based on Variance Weighted Factor Score Distances

State	1	2	3	4	5	6	7
Alabama	.96	-.00	.04	-.02	.00	.05	-.03
South Carolina	.95	-.10	-.01	-.03	.02	.01	-.01
Mississippi	.95	-.06	-.02	.08	-.16	.01	-.06
Georgia	.93	.02	.06	-.03	-.01	.04	.03
North Carolina	.80	-.01	-.04	.06	.23	.06	.05
Arkansas	.78	-.04	-.09	.16	.25	.00	-.05
Tennessee	.71	.03	-.06	.09	.29	.11	.00
Virginia	.57	.09	.03	.02	.13	.11	.21
Louisiana	.57	.20	.26	-.07	-.17	-.10	.07
Texas	.49	.33*	.23	.05	-.05	.05	.15
Florida	.47	.31*	.21	-.01	-.01	.02	.28
Oklahoma	.33*	.06	.30*	.30*	.14	.17	.03
New York	-.03	1.04	-.05	.01	-.00	-.03	-.10
California	-.06	.94	.08	.03	-.04	-.08	-.05
New Jersey	.01	.92	.05	-.04	.07	.06	-.04
Delaware	.04	.77	-.05	.04	-.00	.07	.26
Illinois	.01	.73	-.18	.09	-.04	.30*	.04
Connecticut	-.04	.55	.25	-.03	.13	.32*	.04
Maryland	.23	.53	.11	-.05	.05	.15	.27
Michigan	.08	.51	.15	.03	.04	.37*	.07
Utah	.02	-.01	.77	.16	.14	.03	-.05
Colorado	.02	.12	.68	.20	.02	.13	.06
Wyoming	.02	.01	.67	.32*	-.02	-.10	.13
Oregon	-.01	.00	.63	.23	.15	.23	.05
Washington	-.02	.10	.62	.14	.09	.29	.04
New Mexico	.31*	.10	.59	.06	-.01	-.16	.24
Massachusetts	.00	.24	.50	-.14	.22	.44	-.06
Arizona	.24	.32*	.49	.04	-.07	-.15	.26
Minnesota	.03	.14	.38*	.36*	-.02	.35*	.03
South Dakota	.08	.04	.03	.83	.15	-.07	.04
North Dakota	.07	.03	.11	.83	.11	-.11	.04
Nebraska	.03	.07	.03	.75	-.02	.29	.04
Iowa	.01	.06	.02	.71	.05	.33*	-.01
Montana	.04	-.01	.38*	.55	.14	.02	.12
Idaho	.01	-.00	.16	.53	.45	-.06	.04
Wyoming	.07	.04	.37	.52	.05	.22	.02

TABLE A.6
Continued

State	1	2	3	4	5	6	7
Maine	-.06	.01	.03	.04	.94	-.01	.02
West Virginia	.19	-.04	-.01	-.00	.83	-.07	.00
Vermont	-.06	.03	.07	.20	.81	-.00	.05
New Hampshire	-.09	.02	.20	.01	.78	.20	.06
Kentucky	.41	-.02	-.18	.12	.58	.08	.04
Pennsylvania	.12	.17	.05	.00	.10	.72	.09
Indiana	.11	.10	.01	.18	.06	.71	.15
Ohio	.09	.22	.04	.07	.04	.70	.11
Wisconsin	.02	.11	.02	.24	.03	.69	.14
Missouri	.20	.08	-.07	.21	.08	.65	.12
Rhode Island	.10	.13	.42	-.16	.29	.49	-.04
Alaska	-.05	-.11	-.07	-.01	.03	.05	.69
Nevada	.01	.25	.30*	.09	.03	-.02	.57
Hawaii	.08	.46	.09	.08	-.02	-.01	.54

TABLE A.7
Simple Loadings Factors
Distance Analysis Based on Factor Scores
Weighted According to Percentage of Educational Variables

State	1	2	3	4	5	6	7
South Carolina	.95	-.03	-.04	-.02	.00	.04	.01
Alabama	.93	-.00	-.01	.06	.01	.03	-.05
Georgia	.92	-.01	-.02	.08	.04	.03	.01
Mississippi	.91	.00	.08	-.01	-.00	-.15	-.07
Arkansas	.80	.10	.17	-.08	-.04	.15	-.01
North Carolina	.80	.12	.05	-.03	-.07	.17	.10
Louisiana	.78	-.05	-.08	.15	.32*	-.15	-.02
Tennessee	.78	.20	.08	-.00	-.05	.22	.02
Virginia	.67	.10	.01	.12	.01	.12	.22
Texas	.53	.06	.11	.33*	.19	-.04	.05
Florida	.53	.02	.04	.31*	.19	-.00	.21
Kentucky	.43	.33*	.13	-.14	-.13	.38*	.09
Indiana	.02	.93	.05	-.04	.05	.02	.05
Missouri	.12	.92	.08	-.07	-.01	-.01	.02
Ohio	.01	.90	-.03	.09	.06	.02	.02
Wisconsin	-.06	.87	.14	.00	.04	.02	.05
Pennsylvania	.03	.86	-.10	.09	.06	.11	-.01
Illinois	.04	.56	.08	.44	-.14	-.12	.06
South Dakota	.10	-.04	.86	-.00	.07	.05	.06
North Dakota	.08	-.06	.83	-.03	.16	.03	.06
Nebraska	.03	.23	.79	.11	-.07	-.01	.00
Iowa	-.00	.28	.75	.10	-.08	.05	-.05
Idaho	.03	-.04	.58	-.02	.20	.37*	.07
Wyoming	.05	.11	.58	.15	.25	.12	-.02
Montana	.02	.01	.54	.01	.39*	.15	.12
Minnesota	.00	.22	.46	.27	.21	.08	-.03
Oklahoma	.30*	.14	.31*	.14	.26	.16	-.03
New York	.09	.04	.07	.87	-.07	.00	.01
New Jersey	.09	.07	.02	.82	.01	.11	.02
California	.04	.02	.18	.73	.12	-.06	.05
Delaware	.12	.17	.10	.59	-.04	-.05	.32*
Connecticut	-.03	.25	-.04	.58	.14	.25	.04
Maryland	.29	.12	.01	.52	.05	.07	.24
Michigan	.08	.38*	.08	.50	.06	.08	.03
Massachusetts	-.01	.19	-.10	.46	.28	.45	-.11

TABLE A.7
Continued

State	1	2	3	4	5	6	7
New Mexico	.27	.09	-.03	-.06	.77	-.07	.15
Utah	-.02	.04	.12	.03	.74	.21	-.12
Wyoming	-.01	.04	.28	-.06	.73	-.03	.07
Arizona	.25	.08	.02	.12	.65	-.12	.18
Colorado	-.01	.09	.25	.17	.61	.11	-.00
Oregon	-.05	.08	.28	.16	.51	.29	.01
Washington	-.06	.12	.21	.27	.66	.28	.00
Maine	.03	.02	.11	.01	-.02	.86	.06
New Hampshire	-.05	.18	.02	.07	.00	.77	.07
West Virginia	.26	.02	.04	-.06	.01	.76	.06
Vermont	.01	.02	.27	.02	.04	.72	.09
Rhode Island	.07	.25	-.16	.37*	.21	.50	-.11
Alaska	-.07	.00	-.04	-.05	-.04	.04	.65
Nevada	.04	.03	.12	.18	.35*	.03	.56
Hawaii	.14	.07	.14	.33	.14	-.06	.54

TABLE A.8
Varimax Rotated Factor Loadings
Q Analysis Based on Correlation

State	1	2	3	4	5	6	7	8
Arkansas	.89	-.07	.19	.05	-.12	.18	-.07	-.17
Alabama	.88	-.10	-.22	-.18	-.12	.06	-.02	-.15
South Carolina	.88	-.21	-.23	-.07	-.12	.02	-.01	.06
Tennessee	.87	-.01	-.11	-.02	-.00	.28	-.01	-.11
North Carolina	.86	-.22	-.18	.02	.05	.08	-.12	-.03
Mississippi	.85	-.22	.10	-.14	-.06	.02	-.06	-.09
Washington	-.83	-.27	-.11	.05	.04	.19	-.13	-.20
Georgia	.81	-.06	-.31*	-.20	-.18	.01	.09	-.04
Kentucky	.79	.09	.16	.25	.09	.23	-.05	.01
Oregon	-.75	-.29	.14	.16	-.05	.23	-.16	-.18
Colorado	-.71	-.39*	.11	-.07	-.11	.34*	.18	.10
Virginia	.65	-.25	-.37*	-.11	-.06	.19	.14	.44
Louisiana	.62	-.21	-.25	-.39*	-.25	-.14	-.15	-.07
California	-.62	.13	-.42	-.35*	.15	.12	.29	-.23
Connecticut	-.56	.49	-.44	.15	.07	-.20	-.12	.16
West Virginia	.52	-.16	-.08	.48	.06	.27	-.37*	-.14
New Mexico	-.06	-.79	-.06	-.35*	-.16	-.04	-.08	.12
Illinois	-.35*	.75	-.22	-.17	.31*	-.03	.01	.15
Pennsylvania	-.18	.74	-.33*	.08	.35*	-.06	-.23	-.05
Missouri	.14	.73	.05	.08	.27	.32*	.14	.04
Rhode Island	-.19	.73	-.38*	.23	-.11	-.15	-.08	.01
Utah	-.44	-.71	-.04	-.04	.11	.12	-.18	-.20
Massachusetts	-.43	.70	-.34*	.19	-.14	-.15	-.11	.03
Arizona	-.20	-.67	-.14	-.34*	-.02	-.00	.37*	.10
New York	-.43	.66	-.37*	-.16	-.12	-.04	-.07	.01
Ohio	-.23	.62	-.24	-.00	.60	-.02	-.03	.00
New Jersey	-.45	.62	-.40	-.05	.08	-.18	-.08	.16
Wyoming	-.46	-.56	.50	-.10	-.22	.05	-.10	-.04
Wisconsin	-.29	.56	.27	.10	.41	-.34*	-.25	-.08
Alaska	-.14	-.53	-.26	.03	-.14	-.48	.23	.27
South Dakota	.06	-.20	.92	.09	-.11	.05	-.11	-.08
Nebraska	-.18	.11	.91	.07	.05	.11	.00	.06
Iowa	-.13	.16	.87	.13	.17	.14	-.13	-.03
North Dakota	.04	-.22	.86	.04	-.14	-.01	-.22	-.09
Montana	.30*	-.35*	.74	.04	-.23	.03	-.01	-.04
Kansas	-.32*	-.22	.72	.02	.06	.36*	-.01	.11
Idaho	-.08	-.37*	.65	.24	-.06	.01	.06	-.31*
Maryland	-.18	.25	-.61	-.23	.00	-.20	.03	.54
Minnesota	-.53	.20	.53	-.07	.11	-.10	-.33*	-.28
Maine	.00	.24	.11	.87	.09	-.03	-.01	-.22
Vermont	-.11	.01	.36*	.74	-.03	.14	-.24	.10
New Hampshire	-.26	.50	-.10	.72	.05	-.10	-.02	.04
Texas	.35*	-.11	-.18	-.54	-.00	.41	.28	.06

TABLE A.8
Continued

State	1	2	3	4	5	6	7	8
Indiana	-.06	.15	-.04	.12	.88	.19	-.16	.07
Michigan	-.40	.29	-.42	-.14	.59	-.15	-.20	-.27
Oklahoma	.11	-.27	.34*	-.12	-.05	.71	.11	-.16
Hawaii	-.02	-.11	-.20	-.12	-.21	-.41	.14	-.06
Florida	.06	.02	-.33*	-.26	-.23	.12	.75	-.03
Nevada	-.40	-.33*	-.13	-.10	-.13	-.26	.68	.00
Delaware	-.42	.04	-.34*	-.20	.13	-.15	-.09	.88

BIBLIOGRAPHY

Bailey, J. P. and Guertin, W. H. Test item dependence of several oblique factor solutions. Educational and Psychological Measurement, 1970, 30, 611-19.

Banks, A. S. and Gregg, P. H. Grouping political systems: Q-factor analysis of A cross-polity survey. American Behavioral Scientist, 1965, 9, 3-6.

Berry, B. J. L. A method for deriving multi-factor uniform regions. Polish Geographical Review, 1961, 33, 263-82.

Burt, C. L. Correlations between persons. British Journal of Psychology, 1937, 28, 59-96.

Carlson, H. B. and Harrell, W. Voting groups among leading congressmen obtained by means of the inverted factor technique. Journal of Social Psychology, 1942, 16, 51-61.

Carroll, J. B. Oblimin rotation solution in factor analysis. Computing program for the IBM 704. Cambridge: Harvard University, 1958.

Cattell, R. B. The dimensions of culture patterns by factorization of national characters. Journal of Abnormal and Social Psychology, 1949, 44, 443-69.

Cattell, R. B. The principal culture patterns discoverable in the syntal dimensions of existing nations. Journal of Social Psychology, 1950, 32, 215-53.

Cattell, R. B. Factor analysis: an introduction and manual for the psychologist and social scientist. New York: Harper and Row, 1952.

Cattell, R. B. Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966

Cattell, R. B. and Adelson, M. The dimensions of social change in the U.S.A. as determined by P-technique. Social Forces, 1951, 30, 190-201.

Cattell, R. B., Breul, H., and Hartman, H. P. An attempt at more refined definitions of the cultural dimensions of syntality in modern nations. American Sociological Review, 1951, 17, 408-21.

Cattell, R. B. and Muerle, J. L. The "Maxplane" program for factor rotation to oblique simple structure. Educational and Psychological Measurement, 1960, 20, 269-90.

Chang, H. J. Socioeconomic and educational bases for grouping Florida counties. Paper read at the Southeastern Psychological Association, Miami, 1971.

Cole, J. P. and Smith, G. A. An introduction to factor analysis with the United States as an example. (Bulletin of quantitative data for geographers no. 8.) Nottingham, England: University of Nottingham, 1967.

Cronbach, L. J. and Gleser, C. G. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-73.

Eber, H. W. Toward oblique simple structure: a new version of Cattell's Maxplane rotation program for the 7094. Multivariate Behavioral Research, 1966, 1, 112-25.

Folger, J. K. and Nam, C. B. Education of the American population. A 1960 census monograph. Washington, D.C.: U. S. Government Printing Office, 1967.

Gosnell, H. F. Machine politics: Chicago model. Chicago: University of Chicago Press, 1937.

Gosnell, H. F. and Gill, N. N. A factor analysis of the 1932 presidential vote in Chicago. American Political Science Review, 1935, 29, 967-84.

Gosnell, H. F. and Schmidt, M. J. Factorial and correlational analysis of the 1934 vote in Chicago. American Statistical Association Journal, 1936, 31, 507-18.

Guertin, W. H. The search for recurring patterns among individual profiles. Educational Psychological Measurement, 1966, 26, 151-65.

Guertin, W. H. Typing ships with transpose factor analysis. Educational and Psychological Measurement, 1971, 31, 397-405.

Guertin, W. H. and Bailey, J. P. Introduction to modern factor analysis. Ann Arbor, Michigan: Edwards Brothers, 1970.

Hackney, S. Southern violence. *American Historical Review*, 1969, 74, 99-113.

Harman, H. H. *Modern factor analysis*. Chicago: University of Chicago Press, 1967.

Harris, C. W. A factor analysis of selected Senate roll-calls, 80th Congress. *Educational and Psychological Measurement*, 1948, 8, 583-91.

Hartshorn, T. A. The spatial structure of socioeconomic development in the Southeast, 1950-1960. *Geographic Review*, 1971, 61, 265-83.

Hatt, P., Farr, N. L., and Weinstein, E. Types of population balance. *American Sociological Review*, 1955, 20, 14-21.

Hofferbert, R. I. Socioeconomic dimensions of the American states, 1890-1960. *Midwest Journal of Political Science*, 1968, 12, 401-18.

Hofstaetter, P. R. A factorial study of cultural patterns in the U. S. *Journal of Psychology*, 1951, 32, 99-113.

Jennrich, R. I. and Sampson, P. F. Rotation for simple loadings. *Psychometrika*, 1966, 31, 313-23.

Kaiser, H. F. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, 23, 187-200.

Odum, H. W. *Southern regions of the United States*. Chapel Hill, N. C.: University of North Carolina Press, 1936.

Odum, H. W. and Moore, H. E. *American regionalism: a cultural-historical approach to national integration*. New York: Henry Holt, 1938.

Ogburn, W. F. Factors in the variation of crime among cities. *Journal of American Statistical Association*, 1935, 30, 12-34.

Price, D. O. Factor analysis in the study of metropolitan centers. *Social Forces*, 1942, 20, 449-55.

Rummel, R. J. *Applied factor analysis*. Evanston: Northwestern University Press, 1970.

Russett, B. M. *International regions and the international system*. Chicago: Rand McNally, 1967.

Schutz, R. E. A factor analysis of educational development in the United States. *Educational and Psychological Measurement*, 1956, 16, 324-32.

Smith, D. M. Towards a geography of social well-being. Paper read at annual meeting of the Association of American Geographers, Boston, 1971.

Spearman, C. The proof and measurement of association between two things. American Journal of Psychology, 1904, 15, 72-101.

Stephenson, W. The inverted factor technique. British Journal of Psychology, 1936, 26, 344-61.

Sutton, R. L. State environment as a variable: an essay in concept building. Paper read at Seventh World Congress, International Political Science Association, Brussels, Belgium, 1967.

Thorndike, E. L. Your city. New York: Harcourt, Brace and Company, 1939.

Thurstone, L. L. Multiple factor analysis. Psychology Review, 1931, 38, 406-27.

BIOGRAPHICAL SKETCH

Martha Jean (Roby) Chang was born in Columbus, Ohio, on June 9, 1934. In 1954 she graduated from Malone College with the degree of Bachelor of Religious Education magna cum laude. In 1957 she received a Bachelor of Arts degree summa cum laude from Roberts Wesleyan College with a major in history. In August, 1957, she was married to Richard T. Chang. She attended the University of Rochester on a scholarship and in 1958 received a Master of Education degree. For three years she taught in public schools in New York, Connecticut, and Michigan while her husband pursued graduate work. During the year 1961-62 she attended graduate school at the University of Michigan.

During the years from 1962 to 1969 Mrs. Chang spent most of her time at home with her family. Perry was born in Japan in October, 1962, Penny in Ohio in November, 1963. After her husband received his doctorate and began teaching, home became subsequently Ravenna, Ohio; Lawrence, Kansas; and Pittsburgh, Pennsylvania. In 1966 Mrs. Change received the degree of Master of Arts in history from the University of Michigan. That summer her husband accepted a teaching position in the history department of the University of Florida and the family moved to Gainesville. Mrs. Chang taught junior high school during the year 1966-67. In

1967-68 she helped found the Millhopper Cooperative Nursery School, serving as secretary and member of the executive board. In 1968-69 she was church school superintendent for the United Church of Gainesville. She served on the executive board of the United Protestant Ecumenical Ministry from 1968 to 1971. She is currently a member of the American Educational Research Association.

From September, 1969, through May, 1971, Mrs. Chang was enrolled full time in the College of Education, the University of Florida, during which she was recipient of a Graduate Fellowship. Since October, 1971, she has been employed full time as Research Specialist for PAIS project, P. K. Yonge Laboratory School, the University of Florida.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Wilson M. Guertin, Chairman
Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Alfred B. Clubok
Associate Professor of
Political Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William B. Ware
Assistant Professor of
Education

This dissertation was submitted to the Dean of the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

November, 1971

Dean, College of Education

Dean, Graduate School